AD-A053 120   TRACOR INC AUSTIN TEX   F/G 15/1
ADAPTIVE SPACE PROCESSING STATE OF THE ART: A SUMMARY AND DISCU--ETC(U)
NOV 66 R G BALDWIN
UNCLASSIFIED   TRACOR-66-577-U   NL

1 OF 1
AD
A053120

END
DATE
FILMED
5-78
DDC

Document Number
TRACOR 66-577-U
TRACOR Project 19039

TECHNICAL NOTE

ADAPTIVE SPACE PROCESSING
STATE OF THE ART
A SUMMARY AND DISCUSSION

For

Naval Ship Systems Command

Department of the Navy

Attn: Code 2114

3 November 1966

TRACOR
6500 TRACOR LANE • AUSTIN, TEXAS 78721 • 512 926-2800

TRACOR, INC. 6500 TRACOR LANE, AUSTIN, TEXAS 78721

Document Number
TRACOR 66-577-U
TRACOR Project 19039

9 TECHNICAL NOTE

10 R. G. Baldwin

6 ADAPTIVE SPACE PROCESSING
STATE OF THE ART:
A SUMMARY AND DISCUSSION.

For

Naval Ship Systems Command
Department of the Navy
Attn: Code 2114

11 3 Nov[redacted]966

12 83p.

14 TRACOR-66-577-U

Submitted:

R G Baldwin
R. G. Baldwin

352 100

JOB

TABLE OF CONTENTS

PREFACE

This paper was written to serve as an educational aid for those persons interested in but generally unfamiliar with adaptive space processing.

No attempt has been made to present any new material. Instead, the author has attempted to survey the field of adaptive space processing and to discuss those concepts and techniques deemed to be important.

A consolidated bibliography is provided. The bibliography and the text are equally important, since it is in the items of the bibliography that the serious reader will find rigorous mathematical developments of the concepts and techniques involved.

1. INTRODUCTION

It has become increasingly apparent over the past several years that serious consideration should be given to the use of adaptive space processing for the detection, classification, and localization of enemy submarines. A number of papers have been written which deal extensively with the theoretical aspects of both adaptive and non-adaptive space processing. The bibliography lists several such papers.

Although many papers are available which deal rigorously with the theory, there is a lack of published data relative to the practicality of implementation of adaptive space processing. Little has been said, for example, regarding the computational tradeoffs available for different types of systems operating in different types of signal and noise fields.

The objective in applying adaptive space processing is to design a system which will change its operating characteristics as a function of the spatial organization of signal and noise so as to provide the best processing possible to combat the then existing spatially organized noise field.

For example, many sonar spatial processing systems (or beamformers) are designed under two assumptions, neither of which may be valid in many cases. The first assumption which is commonly made is that the noise field has a spatially isotropic distribution. That is, equal noise energy propagates from all directions in three dimensional space. This assumption is certainly not valid for a passive sonar system mounted in a submarine providing protection for a convoy. For this case, the primary interfering noise may well be generated by the ships in the convoy and as such, would propagate generally from the direction of the convoy. An adaptive spatial processing system would adjust itself so as to suppress the high intensity convoy noise and accept lower intensity noise from some other direction.

A second assumption which is often made is that the interfering noise is describable as plane waves and hence can be analyzed through the use of beam patterns. It is unlikely that ship self noise which couples into the transducer through the structure appears to propagate as plane waves. Nonetheless, an adaptive processor would adjust itself so as to combat this type of noise.

In general the potential effectiveness of adaptive space processing depends on whether or not there is an exploitable difference in the way signals and noise are organized in space.

Work which has made direct contributions to the technology of sophisticated spatial processing is listed below. Undoubtedly, significant contributions have been made by other persons, whose work has not come to the author's attention.

Dr. Finn Bryn, while on a Royal Norwegian Council of Scientific and Industrial Research fellowship, at the Marine Physical Laboratory of the Scripps Institution of Oceanography, developed a theory of optimum signal processing for three-dimensional arrays operating on Gaussian signals and noise. Portions of his work were submitted to "The Journal of the Acoustical Society of America" in July 1961, and published in March 1962 (12).

During the same time period, Dr. Milo Backus, Dr. L. Strickland, Mr. John Burg and associates, all of Texas Instruments Incorporated, working under contract to the U. S. Air Force, developed a theory for spatial processing of seismometer arrays based on the Wiener least-mean-square-error criterion. The final report on this work was submitted to the Air Force in midyear 1961. The work was later published in "Geophysics" in October 1964 (17).

This work was generally concerned with an optimum frequency-domain solution to the array processing problem as was the work done by Bryn. By midyear 1963, Burg had developed and utilized an algorithm for the solution of the Wiener optimum array processing problem completely in the time domain. To the knowledge of the author, Burg has never published the results of this work other than in reports submitted to the Air Force. However, a description of the work is contained in reference (104).

During the same general time period, Dr. Ralph A. Wiggins of the Department of Geology and Geophysics, Massachusetts Institute of Technology, and Dr. Enders A. Robinson, Seismological Institute, Uppsala University, Uppsala, Sweden, became interested in the same problem and published their results (123) in the Journal of Geophysical Research in April, 1965.

Dr. H. Mermoz (7) of the Institute Polytechnique, Grenoble, France, published two papers in 1964 which were concerned with the optimum utilization of an array for separation of a signal of known waveform from noise.

All of the contributions mentioned thus far were concerned with matrix-inversion techniques for the optimum solution to the array processing problem. In May 1965, Capt. S. W. W. Shor of the Anti-Submarine Warfare Systems Project Office, U. S. Department of the Navy, submitted a paper (94) to The Journal of the Acoustical Society of America which was concerned with an iterative solution to the problem of optimum utilization of an array. The paper was subsequently published in January 1966.

Professor Bernard Widrow at Stanford University (102) became interested in the problem in 1965 and in late 1965 he and an associate, Mr. Burwell Goode of NEL, began work on a research program concerned with the utilization of machine learning techniques for the optimum solution of adaptive space processing

problems. As yet their work is unpublished. It is the opinion of the author that Professor Widrow and Mr. Goode have made a significant contribution to the technology of truly adaptive spatial processing.

The bulk of the work which has been done to date has been based on three different optimization criteria:

(1) Maximization of signal-to-noise ratio.

(2) Maximization of likelihood ratio for decision purposes.

(3) Extraction of best estimate of signal waveform.

The first criterion is exemplified by Shor (94), who provides an iterative technique for maximizing the signal-to-noise ratio at a single frequency in the output of an array processor. His work is based on earlier work by Dr. H. Mermoz (7).

Widrow (102) and Goode have extended and simplified the iterative techniques so that they are applicable to either narrow-band or wide-band array processing. Their criterion is to hold the output signal power constant while minimizing the output noise power.

The second criterion is exemplified by Bryn (12) who provides a method for obtaining an output from the array which can be used for statistical decision purposes. Burg (17) provides a method of optimizing on the basis of the third criterion.

Since the techniques presented in these five references are representative of the three basic criteria, the bulk of the present discussion will be based on the five references given above.

This paper presents a discussion of some of the important characteristics of the different techniques including computational requirements.

The reader is referred to the bibliography for detailed and rigorous treatments of the mathematical and statistical concepts involved.

2. SUMMARY AND CONCLUSIONS

The computational requirements for the Wiener, Bryn, Shor, and Widrow adaptive space processing techniques were evaluated relative to three different multiple beam array processing problems. It was concluded, for the three cases considered, that the Bryn matrix inversion technique is computationally more attractive than either the Wiener frequency domain technique or the Shor iterative technique.

For multiple beam problems, this should in general be true since the Bryn technique requires only one noise matrix inversion no matter how many beams are to be formed simultaneously. The Wiener technique requires a separate matrix inversion for each adaptive beam and the Shor technique requires a new iteration for each adaptive beam.

It was further concluded that the Widrow technique was more attractive than the Bryn technique in terms of storage requirements. No conclusions could be drawn relative to the computational difficulty of the Widrow technique since the convergence rates are presently unavailable.

3. CONCEPTS AND OPTIMIZATION CRITERIA

The following paragraphs present a discussion of concepts and optimization criteria which exemplify the three criteria presented in the introduction.

3.1 WIENER CRITERION

The Texas Instruments approach on adaptive beamforming, as outlined in Burg's paper and in the NEL report(104), makes use of the Wiener least-mean-square error criterion. The technique can be applied in either the time domain or the frequency domain. In either case, the design is such that an attempt is made to extract the signal from the noise in an undistorted fashion.

That is, the processor operates on a set of transducer outputs, each containing signal-plus-noise, and attempts to provide an output consisting of a noise free reproduction of the signal as it appears on a reference transducer. It is generally not possible to produce a noise free output. In that event, the output noise level is reduced as much as possible based on any differences that exist between the organization of signals and noise in space. The output signal is a near reproduction of the reference signal added to the output noise. With this technique, signal distortion, as well as output noise, contributes to the error term. Hence, signal distortion is minimized.

The differences in the spatial organization need not necessarily be simple differences in the direction of propagation. For example, if the ship self noise couples into the different transducers at different levels but in such a manner that the noise output from one transducer is strongly correlated with the noise output from another transducer, then this represents a type of space organized noise which can be handled with any of the techniques under discussion, but which is not generally describable by use of beam patterns.

One important characteristic of the Wiener optimization criterion is that knowledge of the signal waveform is not required. Shor indicates that his technique requires knowledge of the exact waveform of the expected signal. It is not clear why, unless he requires this simply to satisfy the requirements of the matched filter which he cascades with the output of the adaptive space processor. In general, it is the relative phase characteristic between hydrophones which is important in space processing and not the absolute phase characteristics.

The information required for implementation of the Wiener technique is the cross-correlation sets describing signal and noise. The cross-correlation set describing signal can probably be synthetically generated once the characteristics of the array elements are well in hand. The cross-correlation set describing noise, however, must be repeatedly measured for an adaptive system. It is important to note that if it is desired to form a set of adaptive beams simultaneously, using the same set of transducers, then only one set of noise correlation functions is required to design the full set of beam programs. It is necessary, however, to either measure or otherwise provide a signal correlation set for each adaptive beam design.

3.2 SHOR CRITERION

The Shor criterion is very straightforward. That is, design a spatial processing system which will maximize the signal-to-noise ratio in the output. He shows that the output signal-to-noise ratio will be driven to an extreme value whenever

$$\frac{\overline{S_i S_o}}{\overline{S_o S_o}} = \frac{\overline{N_i N_o}}{\overline{N_o N_o}} \quad \text{for all } i \text{ ,}$$

where $\overline{S_i S_o}$ is the cross correlation between the output signal and the input signal on channel i,

$\overline{N_i N_o}$ is the cross-correlation between the output noise and the input noise on channel i, and

$\overline{S_o S_o}$ and $\overline{N_o N_o}$ are output signal power and noise power respectively.

For purposes of comparison, the Wiener criterion is satisfied whenever

$$\overline{S_i S_o} + \overline{N_i N_o} = \overline{S_i S_r} \text{ for all } i \ ,$$

where $\overline{S_i S_r}$ is the signal cross-correlation between each input channel and a reference channel.

The Bryn criterion, to be discussed later, is satisfied whenever

$$\overline{N_i N_o} = \overline{S_i S_r} \ .$$

It is interesting to note that at least for unidirectional signal models, a set of optimum answers obtained with either the Wiener or Bryn techniques will satisfy the Shor criterion. At least, the author has been unable to show that this is not true and can demonstrate that it is true for numerous simple problems.

There is one case, that of perfectly coherent noise, for which the Shor technique will fail to provide a set of optimum answers. In that event, the normalized noise correlation will approach an indeterminate value. However, since this case is extremely unlikely to occur in nature, this does not present a serious obstacle.

3.3 BRYN CRITERION

The Bryn criterion does not lend itself to nearly so simple a physical interpretation as do the Shor and Wiener criteria. Bryn bases his case on the presentation of a likelihood ratio which can be compared with a known threshold for decision purposes. The argument depends heavily upon Gaussian distributions in order that the final result will be interpretable as a decision making function. In spite of the Gaussian assumptions, Bryn develops a spatial processor which appears to provide the same output signal-to-noise ratio as the Wiener technique.

Implementation of the technique involves the inverting of a matrix, the same size as with the Wiener technique. The matrix, however, does not contain the same information. The Bryn technique requires the full cross-correlation matrix for noise, and one column of the cross-correlation matrix for signal. This is less information than is required for the Wiener technique.

The Bryn technique, like the Shor technique, will fail to provide a set of optimum answers whenever the noise is totally coherent. For this case, the Bryn matrix will be singular.

3.4 WIDROW-GOODE CRITERION

Widrow and Goode have been engaged in research related to the use of iterative learning techniques for adaptive processing of hydrophone arrays.

The criterion which they have used can be considered as a signal-to-noise power ratio maximization criterion (102, pg 5). The system operates under two separate (in time) adaptation modes. During the first mode, the system is "taught" to faithfully reproduce signals according to a specific definition of signal. During the second mode, the system is taught to reject noise as received by the hydrophone array. The specific result of the teaching process is a set of weighting factors. These weights

are applied to the hydrophone outputs (either in raw or time delayed form) in a multiply and add operation for the purpose of providing a single array output which has the desired characteristics. As expressed by Goode (121), "With proper switching between the two modes, the signal power is maintained near a constant level, and the noise power is reduced to the minimum possible for the prescribed level of signal power." It should be pointed out that an attempt is made to hold signal distortion to a minimum.

The information required for implementation of the technique is considerably less than is required for implementation of either the Wiener, Bryn, or Shor techniques. Specifically, no cross-power functions are required and hence the need for long data history storage is eliminated. What is required for each iteration is a measurement of the instantaneous values of the multichannel input (signal or noise) and the values of the weights calculated in the previous iteration. The history data are effectively contained in the values of the weights themselves.

Widrow and Goode are also engaged in research related to an iterative processor which simultaneously adapts on signal and noise, thus eliminating the need for two-mode operation. This is analogous to the Wiener signal extraction technique in that the system minimizes the square of the difference between a simulated signal and the output of the processor (Ref. 10, June 1966). Goode mentions that "An analytical and computational study will be necessary to arrive at an accurate description of the operating characteristics of such a system. However, preliminary results indicate a smaller misadjustment is attained using the simultaneous adaptation algorithm, than with the previous algorithm, especially when there is a strong signal in the beam."

4.0 MATHEMATICAL FORMULATIONS

This section contains the mathematical formulation of the design equations for each of the adaptive space processing techniques under discussion. The reader is referred to the bibliography for derivations of the equations by the original authors.

4.1 BRYN TECHNIQUE

Bryn (12) derives the design equations for the system shown in Fig. 4.1. The equations describing the filters are given by

$$Z_i(n) = [U_i^2(n) + V_i^2(n)]^{1/2}\, e^{j\varphi_i(n)} ,$$

where

$$U_i(n) = \sum_{h=1}^{K} [r_{hi}(n) \cos\omega_n\tau_h + r_{h,k+i}(n) \sin\omega_n\tau_h],$$

$$V_i(n) = \sum_{h=1}^{K} [r_{h,k+i}(n) \cos\omega_n\tau_h - r_{hi}(n) \sin\omega_n\tau_h],$$

$$\varphi_i(n) = \tan^{-1} [\frac{V_{i(n)}}{U_{i(n)}}].$$

These equations are formulated in the frequency domain. The index n is a frequency index. The indices h and i are related to the columns and rows of correlation matrices. The number of hydrophones is given by K. The terms $r_{hi}(n)$ and $r_{h,k+i(n)}$ are elements in the matrix which is the reciprocal of a noise correlation matrix Q_n given by

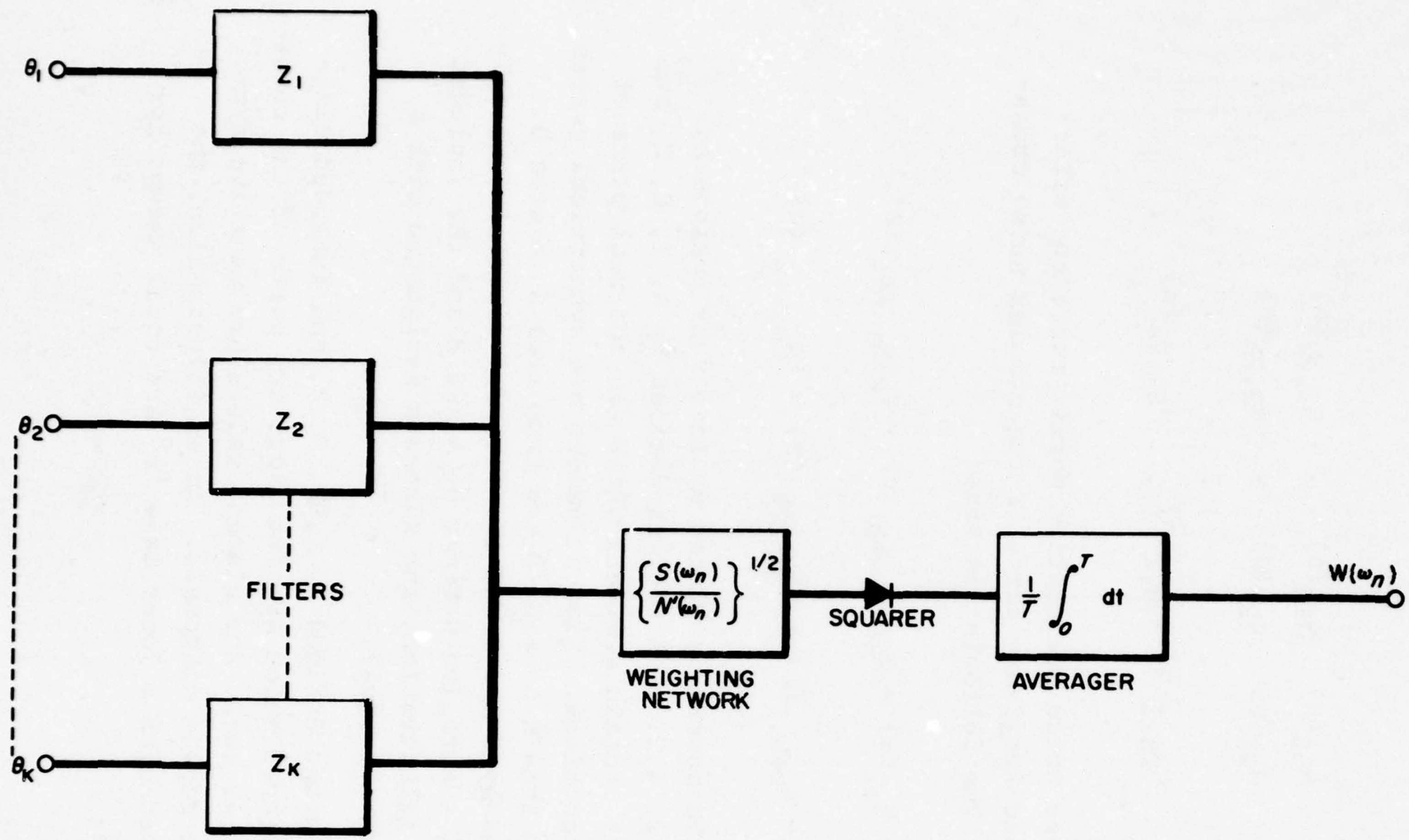

FIG. 4.1 - IMPLEMENTATION OF BRYN PROCESS (FROM BRYN, J. ACOUST. SOC. AMER., MARCH, 1962)

TRACOR, INC. DWG. A6-62-56
AUSTIN, TEXAS 10-26-66 BALDWIN/SL

$$[Q_n] = \begin{bmatrix} q_{11}(n) & q_{12}(n) & \cdots & q_{1,2k}(n) \\ q_{21}(n) & q_{22}(n) & \cdots & q_{2,2k}(n) \\ \cdots & \cdots & & \cdots \\ q_{2k,1}(n) & q_{2k,2}(n) & \cdots & q_{2k,2k}(n) \end{bmatrix} .$$

The elements in the noise correlation matrix represent effectively the real and imaginary parts of the complex noise cross-power functions. The following is true:

$$q_{ih}(n) = q_{hi}(n) = q_{k+i,\ k+h}(n) = q_{k+h,\ k+i}(n),$$

and $$q_{i,k+h}(n) = q_{k+h,i}(n) = -q_{k+i,h}(n) = -q_{h,k+i}(n).$$

In other words, the matrix is broken up into four regions as illustrated by Fig. 4.2. The regions defined by A, B, C, D, and included diagonals contain elements which are the real parts of the cross-power functions. Element values are symmetrical across the diagonal from area A to area B and from area C to area D. The region defined by C and D and the included diagonal is simply a reproduction of the region defined by A and B and the included diagonal. As an illustration, the elements designated with a lower case "a" are all equal.

The regions defined by E, F, G, H, and included diagonals contain elements which are the imaginary parts of the cross-power functions. As such, the element values are negative symmetric across the region diagonals. As an illustration, the elements designated with a lower case "b" are equal except for sign as indicated.

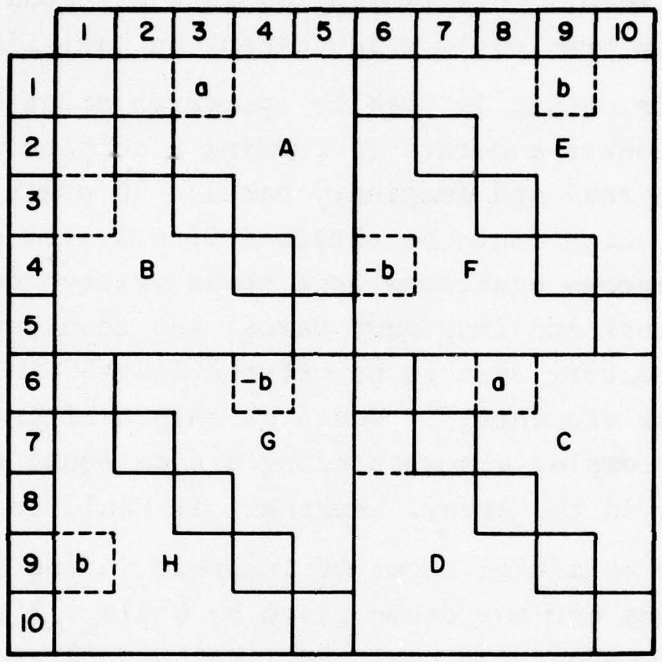

FIG. 4.2 – ILLUSTRATION OF THE BRYN MATRIX

TRACOR, INC. DWG. A6-62-55
AUSTIN, TEXAS 10-26-66 BALDWIN/SL

Although it isn't explicitly indicated by Bryn, the diagonal elements in the regions containing the imaginary part of the cross power must be zero. This is because these diagonal elements represent the imaginary portion of autopower functions which must be zero. This is implied in the second of the two equations given previously which cannot be satisfied for i=h.

This matrix is totally symmetric about the main diagonal and represents a method of forming a complex matrix by separating the real and imaginary parts. In other words, this is the matrix which would be obtained if a system of complex linear simultaneous equations were to be written out so as to separate the real and imaginary parts, and then put into matrix form. If the matrix were to be written instead in a form containing complex elements, it would be only half as large, with the number of complex elements along a side equal to the number of hydrophones in the array. Further, it would be Hermetian.

The remaining terms of interest in the Bryn filter design equations are the terms given by $\mathrm{Cos}(\omega_n \tau_h)$ and $\mathrm{Sin}(\omega_n \tau_h)$. These terms represent the real and imaginary parts, respectively, of the cross power describing a single direction plane wave signal. In this representation, ω_n is the frequency of interest and τ_h is the time delay required to project the hydrophone outputs onto a plane at the rear of the array for a plane wave signal from a single direction.

Since the system design is performed at a single frequency, the output from each hydrophone for a signal can be considered as a sinusoid. The phase of the sinusoid relative to an arbitrary reference is equal to the product of the frequency in radians/second and the time delay in seconds. The cross power between each hydrophone and a hydrophone at the reference can then be written as (A+jB) where $\sqrt{A^2+B^2}$ equal unity,

$$A = \text{Cos}\ \omega_n \tau_h,$$

and

$$B = \text{Sin}\ \omega_n\ \tau_h.$$

Having obtained the physical interpretation of the terms in the Bryn design equations, it is possible to write the equations in another, possibly more conventional form, as shown by the complex matrix given below. It can be shown by expansion that this matrix equation is equivalent to the Bryn design equations.

$$\begin{bmatrix} R_{11}+jI_{11} & R_{12}+jI_{12} \\ R_{21}+jI_{21} & R_{22}+jI_{22} \end{bmatrix} \begin{Bmatrix} U_1+jV_1 \\ U_2+jV_2 \end{Bmatrix} = \begin{Bmatrix} \text{Cos}\ \omega\ \tau_1 - j\ \text{Sin}\ \omega\ \tau_1 \\ \text{Cos}\ \omega\ \tau_2 - j\ \text{Sin}\ \omega\ \tau_2 \end{Bmatrix}$$

(Throughout the text, brackets [ ] are used to denote square matrices and braces { } are used to denote column matrices.) This formulation illustrates that the Bryn design equations can be written in the form

$$\sum_{h=1}^{K} N_{ih}(f) \cdot H_h(f) = S^*_{ir},$$

$$i = 1,K,$$

where $N_{ih}(f)$ is the cross power set describing noise,

$H_h(f)$ is the set of complex filters to be designed,

S^*_{ir} is the signal cross power between each hydrophone and a reference behind the array.

The significance of this formulation of the Bryn design equations is the similarity which it bears to the Wiener formulation to be presented later.

These equations can be written in matrix form as

$$\{H_h(f)\} = [N_{ih}(f)]^{-1} \{S^*_{ir}(f)\} .$$

4.2 MERMOZ TECHNIQUE

Mermoz (7) provides a frequency domain solution for a spatial processing system designed to maximize the signal-to-noise ratio in the output of an array of hydrophones as

$$\sum_k C_{jk}(f) \cdot h_k(f) = Ke^{-i2\pi ft_o} S^*_j(f),$$

where $C_{jk}(f)$ are frequency domain cross-correlation coefficients for the noise field,

$H_k(f)$ are the complex frequency filters required to maximize the signal-to-noise ratio,

$S^*_j(f)$ are the complex conjugates of the Fourier transforms of the actual signal outputs from each hydrophone,

$e^{-i2\pi ft_o}$ is a phase shift associated with an arbitrary definition of zero time,

K is a scaler associated with the signal amplitude.

Since the value of t_o is rather arbitrary, the equations can be rewritten in the following form and considered to be properly defined within an arbitrary time delay.

$$\sum_k C_{jk}(f)\, h_k(f) = K\, S_j^*\, (f),$$

From j=1 to j=N.

This formulation looks very similar to the Bryn formulation. There is, however, one significant difference. In the Mermoz formulation, the terms in the right hand side contain the absolute phase characteristics of the signals in addition to the relative phase characteristics between signals at the output of the various transducers. Hence, the Mermoz formulation will effectively provide a set of filters which, in addition to separating signals from noise on a spatial basis, will also maximize the signal-to-noise ratio by providing "matched" filtering.

4.3 SHOR TECHNIQUE

The Shor processor is shown in Fig. 4.3. The arrangement of the system provides for a complex weighting factor to be applied to each hydrophone output. The imaginary part of the complex weighting factor is obtained by phase shifting the hydrophone outputs by -90^{o} and then applying a weighting factor b_m to the phase shifted output.

The Shor technique is an iterative process whereby the signal-to-noise ratio in the space processor output is maximized by proper selection of the a_m and b_m. The space processor is followed by a conventional matched filter.

The values of the a_m and the b_m are determined by an iterative process which simultaneously minimizes a set of error functions, two for each input channel. The error functions are defined as

$$G_{ai} = \frac{\overline{S_i S_o}}{\overline{S_o S_o}} - \frac{\overline{N_i N_o}}{\overline{N_o N_o}},$$

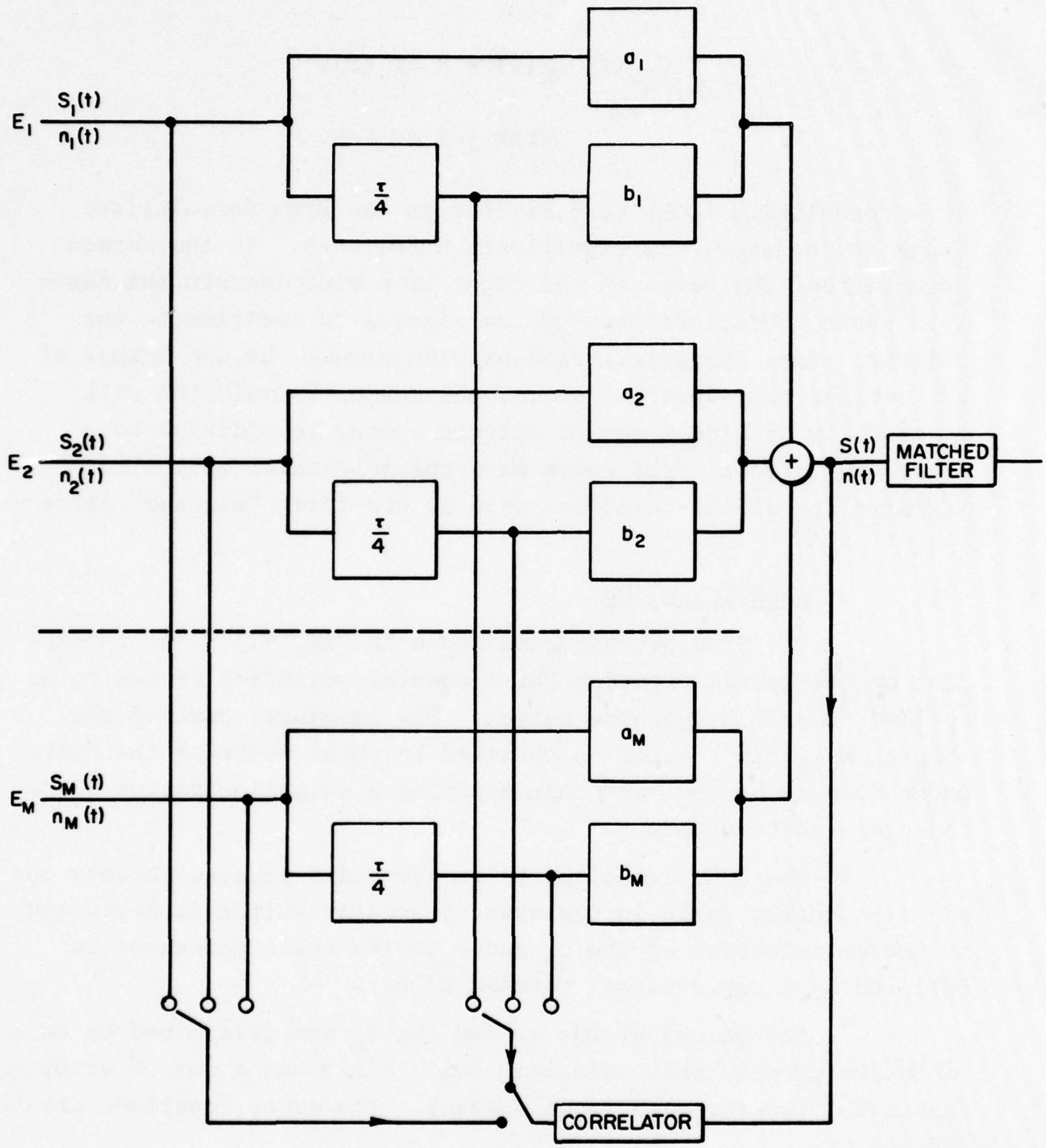

FIG. 4.3 — SHOR ADAPTIVE SPACE PROCESSOR (FROM SHOR, J. ACOUST. SOC. AMER., JAN., 1966)

TRACOR, INC. DWG. A6-62-65
AUSTIN, TEXAS 10·26·66 BALDWIN/SL

and

$$G_{bi} = \frac{\overline{S_i(-T)S_o}}{\overline{S_oS_o}} - \frac{\overline{N_i(-T)N_o}}{\overline{N_oN_o}} ,$$

where $\overline{S_iS_o}$ is the cross correlation at a single frequency between the output signal and the input signal on channel i,

$\overline{S_i(-T)S_o}$ is the cross correlation of a single frequency between the output signal and the time delayed input signal on channel i,

$\overline{S_oS_o}$ is the average power at a single frequency for signal in the output,

$\overline{N_iN_o}$ is the cross correlation at a single frequency between the output noise and the input noise on channel i,

$\overline{N_i(-T)N_o}$ is the cross correlation at a single frequency between the output noise and the time delayed input noise on channel i,

$\overline{N_oN_o}$ is the average noise power at a single frequency in the output.

These error functions are minimized by iteratively evaluating

$$a_i^{(\ell+1)} = a_i^{\ell} + \lambda\, G_{ai}^{(\ell)} ,$$

and

$$b_i^{(\ell+1)} = b_i^{\ell} + \lambda\, G_{bi}^{(\ell)},$$

where $a_i^{(\ell+1)}$ and $b_i^{(\ell+1)}$ are the new values of the weighting coefficients,

a_i^{ℓ} and b_i^{ℓ} are the values of the weighting coefficients from the previous iteration,

λ is a small positive real constant,

$G_{ai}^{(\ell)}$ and $G_{bi}^{(\ell)}$ are the measured errors for the previous iteration.

This technique could be considered for broadband operation by replacing each of the phase shift networks and the a_m and b_m weighting factors by complex frequency filters. Therefore, even though the system is described for a narrow band case, the concept is not limited to narrow band operation. For wide band operation, it would be necessary to measure correlation coefficients at many frequencies.

4.4 WIENER FREQUENCY DOMAIN TECHNIQUE

Burg (17) gives the frequency domain design equations for the Wiener technique as

$$\sum_{j=1}^{N} [S_{mj}(f) + N_{mj}(f)]\, H_j^*(f) = S_{mo}(f),$$

$$m = 1 \text{ to } N,$$

where $S_{mj}(f)$ are the cross power functions describing signal,

$N_{mj}(f)$ are the cross power functions describing noise,

$H_j^*(f)$ are the complex frequency filters to be designed,

$S_{mo}(f)$ are the cross power functions relating the signal on each channel to the signal on a reference channel.

The similarity between this formulation and the Bryn formulation should be noted.

When written in matrix form, the Wiener equations become

$$\{H_j^* (f)\} = [S_{mj}(f) + N_{mj}(f)]^{-1} \{S_{mo}(f)\} .$$

In the Wiener formulation, the matrix to be inverted contains the full cross power set which describes the signal. However, in the Bryn formulation, only the cross power functions describing noise are contained in the matrix to be inverted.

4.5 WIENER TIME DOMAIN TECHNIQUE

The Wiener problem can be formulated in the time domain as well as in the frequency domain. However, it has been indicated by Texas Instruments (104) that the frequency domain design is considerably faster (computationally) than the time domain design. Consequently, the time domain formulation is not considered in detail in this paper.

4.6 WIDROW-GOODE TECHNIQUE

The Widrow-Goode technique has been developed so as to be applicable to both narrow-band and wide-band systems. The implementation concept for narrow-band systems is shown in Fig. 4.4 and that for wide-band systems is shown in Fig. 4.5.

Consider first the narrow-band implementation shown in Fig. 4.4. The system is shown switched into mode II, where the hydrophone outputs are applied to the processor inputs as shown. In this mode, it is desired that the system output power be zero. Therefore, any output is considered to be an error. This error is multiplied by each hydrophone output and each time-delayed hydrophone output. The product so obtained is used to modify the weighting factors, w_i, in a manner to be described in the following paragraphs. While operating in this mode, the system is "taught" to reject noise.

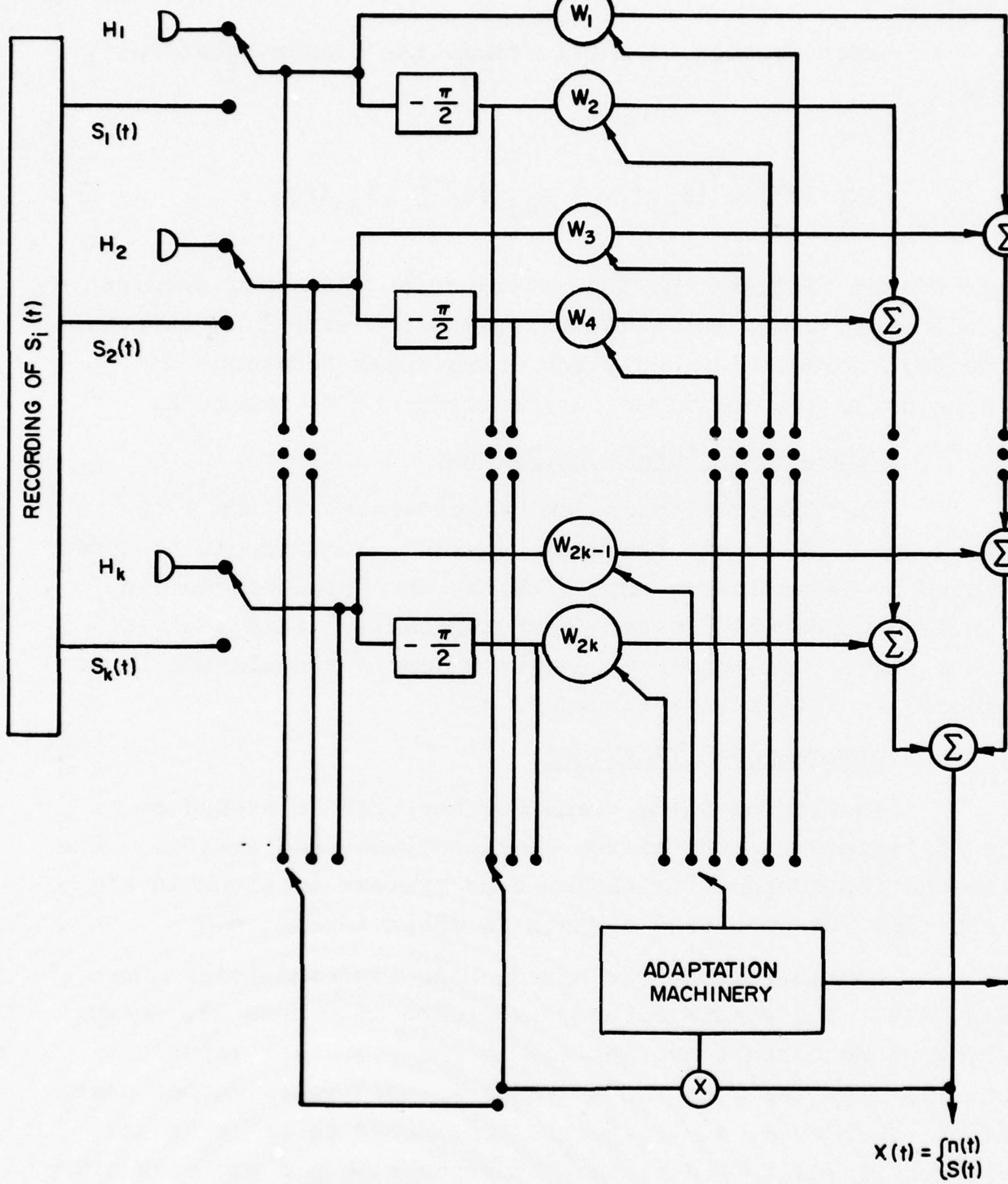

FIG. 4.4 — ADAPTIVE PHASE COMPENSATOR (FROM WIDROW, "PROPOSAL FOR RESEARCH ON ADAPTIVE HYDROPHONE ARRAYS")

TRACOR, INC. DWG. A6-62-57
AUSTIN, TEXAS 10·26·66 BALDWIN/SL

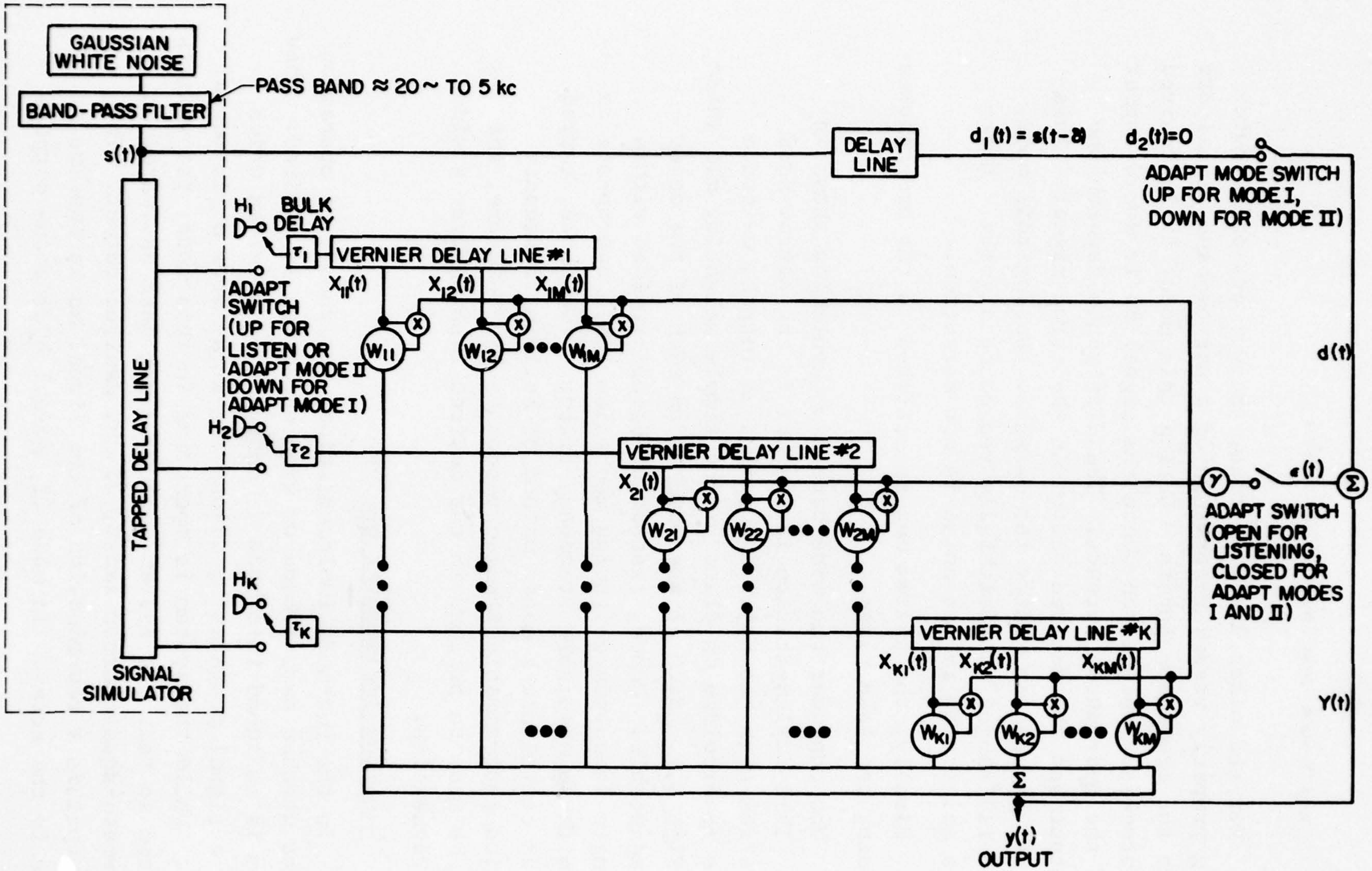

FIG. 4.5– ADAPTIVE PROCESSOR FOR A PASSIVE HYDROPHONE ARRAY, (FROM WIDROW, "PROPOSAL FOR RESEARCH ON ADAPTIVE HYDROPHONE ARRAYS")

TRACOR, INC. DWG. A6-62-58
AUSTIN, TEXAS 10·26·66 BALDWIN/SL

For the other learning mode, the outputs of a device containing properly phased recordings of a narrow-band signal are applied to the processor inputs. During this mode, it is desired that the processor output reproduce the signal as it would appear on one of the hydrophone outputs. The difference between the actual output and the desired output is the error signal. This error signal is used to modify the weights. During this mode, the system is "taught" to faithfully reproduce signals. Hopefully, the ability to reject noise is not forgotten.

Finally, the system can be switched to the hydrophones and listening can take place.

Now consider the wide-band implementation shown in Fig. 4.5. This implementation is similar to the narrow-band implementation with one major exception. A tapped, weighted delay line is provided in place of the simple weighting mechanism shown in Fig. 4.4. Each of the weights on each of the delay lines is adaptable. Hence, instead of being provided with a simple single frequency weighting mechanism, each hydrophone is provided with an adaptable frequency filter. (Weighted, tapped delay lines constitute a handy method of building frequency filters with programmable impulse responses.) Therefore, the array can be made to perform in the desired fashion over a wide band of frequencies.

4.6.1 Two-Mode Implementation

In the two-mode implementation, the theory of operation is the same whether narrow-band or wide operation is desired. When the system is switched into mode I, an input is provided which simulates a signal propagating across the array from a particular direction. While the system is operating in this mode, the weights are adjusted so that the system output is the best estimate, in the least-mean-square-error sense, of the desired output. The desired output is a reproduction of the signal as it appears at some point in the array. In mode II, actual hydrophone outputs

are applied to the system and the weights are adjusted toward zero output.

An expression can be readily derived for the mean-square-error in the signal estimation. The expression for the mean-square-error contains cross power terms, as before. Thus far, little has been said that is different from those techniques discussed previously. The primary difference between the Widrow technique and previously discussed techniques is Widrow's assumption that the instantaneous square error is a good approximation to the mean-square-error. Experimental results have indicated that the assumption appears reasonable. This is a particularly significant approximation since it eliminates the need for the long data history storage requirements inherent in time-averaging type processors.

Widrow's least-mean-square (LMS) adaptation algorithm is expressed in a continuous form as (102)

$$\frac{d\vec{W}}{dt} = \gamma\ \epsilon(t)\ \dot{\vec{X}}(t).$$

The quantity $\vec{W}$ represents a vector defined by the weights, γ is a constant which controls the rate of adaptation, $\vec{X}(t)$ is a vector time function representing the signals as applied to the weights w_{-j} of Fig. 4.5, and $\epsilon(t)$ is a scalar time function which is the difference between the desired instantaneous response $d(t)$ and the actual instantaneous response $y(t)$:

$$\epsilon(t) = d(t) - y(t) = d(t) - \sum_{1}^{M} \sum_{1}^{K} w_{ij}\ x_{ij}(t).$$

The algorithm is derived by obtaining the gradient of the squared error as follows:

$$[\varepsilon(t)]^2 = [d(t) - \sum_1^M \sum_1^K w_{ij}\, x_{ij}(t)]^2 .$$

The change in the squared error which occurs with a change in an individual weight is given by

$$\frac{\partial[\varepsilon(t)]^2}{\partial w_{ij}} = -\,2[d(t) - \sum_1^M \sum_1^K w_{ij}\, x_{ij}(t)]\, x_{ij}(t),$$

which by substitution gives

$$\frac{\partial[\varepsilon(t)]^2}{\partial w_{ij}} = -\,2\varepsilon(t)\, x_{ij}(t).$$

Thus, the gradient of the squared error is seen to be proportional to the product of the instantaneous error and the signal or noise vector $x_{ij}(t)$. By the method of steepest descent, the change in w_{ij} which will reduce the squared error the fastest is a change in the direction of the gradient. Therefore, the iteration algorithm may be written as

$$w_{ij}^{(\ell+1)} = w_{ij}^{(\ell)} + \lambda[\frac{\partial[\varepsilon(t)]^2}{\partial w_{ij}} = w_{ij}^{(\ell)} - 2\lambda\, \varepsilon(t)\, x_{ij}(t),$$

where λ is a positive real constant governing the rate of convergence and the stability.

The iterative technique requires, therefore, that the hydrophone (or delay line tap) outputs, $x_{ij}(t)$ and the error

between the desired output and the actual output, $\epsilon(t)$, be measured at a given time. The weights are then adjusted by calculating the product

$$2\lambda \, \epsilon(t) \, x_{ij}(t).$$

This process can be thought of as searching for a multidimensional vector specified by the set of weights which is as near as possible to being orthogonal to the vectors specified by the various propagating noise components. However, a constraint is placed on the weight vector such that the product of the weight vector and the signal vector yields some desired result. This is discussed in more detail in Appendix A.

4.6.2 Simultaneous Signal and Noise Adaptation

Widrow and Goode are also doing research relative to a system which does not require two adaptive modes. This system simultaneously adapts on both signal and noise to minimize the square of the difference between a simulated signal and the actual processor output. As of this writing, a digital technique had been formulated and reported. However, no reports were available to the author containing experimental results obtained with the technique. The mathematical formulation of the technique as reported by Goode (Ref. 121, June 1966) is quoted below.

"The hydrophones are each sampled at a rate greater than $1/2f_u$, where f_u is the highest frequency passed by the hydrophone filters. (The optimum sampling rate is yet to be determined. The number of bits per sample is also an interesting optimization problem.) The samples are introduced into shift registers. These samples will be denoted by $x_i(t_j)$, the j^{th} sample of the i^{th} hydrophone. Another set of shift registers carry the simulated signal samples $s_i(t_j)$. Each shift register is tapped; that is, several parts of the waveform stored in the

shift register can be used simultaneously. The output of each tap is fed to a multiplier where it is multiplied by a 'weight', the product then being added to all the other products to form a beam. If a hydrophone contributes to more than one beam, each tap is fed to as many multipliers as there are beams using that hydrophone. There need not be as many multipliers as there are taps, for the multipliers may be multiplexed and operated at a faster rate than the shift register clock rate. For every beam, there are two outputs formed using the same set of weights. One output, denoted by X(t), is formed from the data actually received by the hydrophones:

$$X(t) = \sum_{i=1}^{K} \sum_{j=1}^{L} W_{ij} \ x_i(t - t_j),$$

where K = the number of hydrophones or clusters of hydrophones used to form a beam; L = the number of taps on each shift register, and W_{ij} = weights. Another output S(t), distinct from X(t), is formed by using the same weights, W_{ij}, and the simulated signal from the direction of the beam:

$$S(t) = \sum_{i=1}^{K} \sum_{j=1}^{L} W_{ij} \ s_i(t - t_j).$$

Thus, reception may continue without interruption during adaptation. It is seen that the system must be digital to accomplish this objective.

"The desired output is taken to be $s_o(t - \frac{t_L}{2})$. The error $\epsilon(t)$ is defined by

$$\epsilon(t) = X(t) + S(t) - s_o(t - \frac{t_L}{2}).$$

The adaptation algorithm is

$$\Delta W_{ij} = \lambda\, \varepsilon(t)\, [x_i(t - t_j) + s_i(t - t_j)]."$$

5. IMPORTANT CHARACTERISTICS

Even though the techniques under discussion appear capable of providing the same output signal-to-noise ratio, they differ considerably in other respects. This section contains a discussion of some of the ways in which they differ.

5.1 OUTPUT SIGNAL LEVEL

Presumably, any adaptive beamforming technique which is put into operation will provide output signals to other electronic equipment. Consequently, it would be desirable to state with assurance that a signal producing some known output voltage at the transducer would produce some other known voltage at the output of the adaptive space processor. This should be true regardless of the present state of the spatial organization of the background noise field. This is one major area of difference between the adaptive beamforming techniques under discussion.

An important characteristic of any adaptive beamforming technique, therefore, is the total system response for signals on the main beam axis. The Wiener technique and the Widrow technique approximately stabilize the signal response and allow the noise response to vary with the state of the spatial organization of the noise. This is a result of optimizing on a criterion which attempts to extract the signal on a noise-free, distortion-free basis.

On the other hand, with the Bryn technique, both the signal response and the noise response depend on the state of the spatial organization of the noise field. There is nothing in the design equations to constrain the signal response. Hence, with an operational adaptive system, it could be exceedingly difficult to predict with assurance the output voltage level to expect for a given input signal level.

If the output were passed through an AGC, the absolute value of the output signal level would become meaningless. However, in the event that the absolute level of the output signal is important, a simple case has been worked out to illustrate the manner in which the Wiener and Bryn techniques differ in this respect. The results are given in Appendix B.

As is always the case, the Wiener technique does not provide the advantage of a stablized signal output free of charge. The stabilization is purchased at the expense of providing a full signal correlation matrix, which is not needed with the Bryn technique.

Although the Shor technique does not necessarily provide a stabilized signal response, the iterative procedure could be modified to provide such stabilization because, although the answers from the Shor technique are optimum, they are not unique.

5.2 COMPUTATIONAL REQUIREMENTS

There are important differences in the computational requirements for the different techniques. A technique which appears to be computationally most attractive for one problem will not necessarily be most attractive for another.

The spatial processing portion of all the techniques can be illustrated as shown in Fig. 5.1.

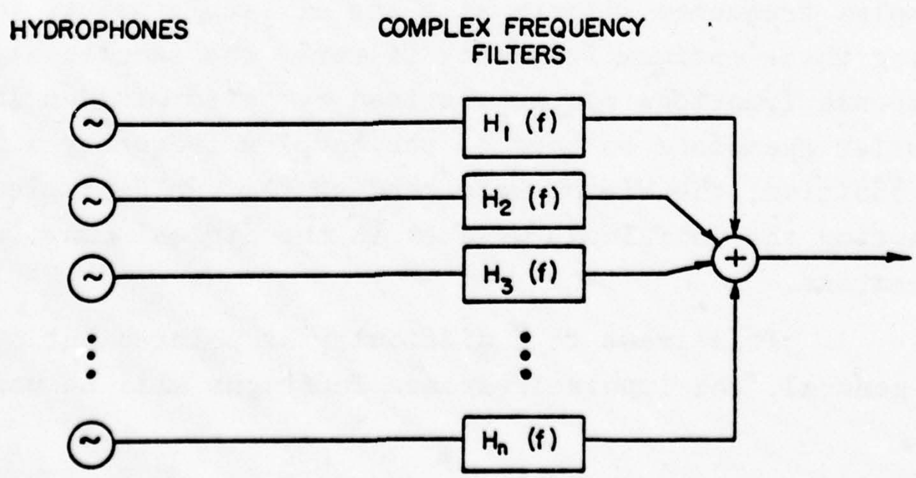

FIG.5.1 - ILLUSTRATION OF SPATIAL PROCESSORS

No matter which design technique is employed, if wide frequency band operation is desired, the spatial processor takes the form of an array of N hydrophones each feeding a complex frequency filter, with the outputs from the filters summed. A matched filter would follow the summing junction with the Shor system and a frequency weighted square and integrate device would follow the summing junction with the Bryn system. Any conventional temporal processing device could follow the summing junction with the Wiener or the Widrow systems.

In general, the frequency filters will be complex with very complicated response characteristics. Further, the response characteristics must be programmable as a function of time. The only method known to the author of obtaining programmable, complicated, complex frequency filters is through the use of shaped impulse response functions. Although this can be done using weighted, tapped analog delay lines, the need for rapid changes in the program makes the use of programmable digital convolution devices more attractive.

The problem, then, is reduced to one of obtaining a set of digital convolution operators, or sampled impulse response functions, which have frequency response characteristics as desired. All the techniques can be implemented by evaluating the design equations at many closely spaced frequencies and then sorting out the answers thus obtained to provide the optimum complex frequency filters at a set of sample points in frequency. Using these optimum frequency filters, the sampled impulse response functions can be obtained by performing an inverse Fourier transform on each of the complex frequency filter functions. In addition, the Widrow wide-band system can be implemented by adapting the individual weights in the digital convolution operators.

This leads to a difficulty as pointed out by Goode (43). In general, the impulse response functions will be non-negative

for negative time. This can be partially compensated for by introducing a finite time delay into the processor and shifting the negative time portion of the response into positive time. The effectiveness of the compensation will depend on how rapidly the impulse response converges to zero for negative time. In any event, this problem serves to illustrate a larger overall problem. That problem is, all of the techniques under discussion can be used to solve for the optimum frequency filters to be applied to the transducers to effect the desired result. However, once having obtained the required frequency filtering characteristics, these filters must be approximately implemented through the use of finite length truncated impulse response functions. The degree of system degradation which results from this truncation is dependent upon the complexity of the noise and signal fields, the required bandwidth, etc.

One major difference between the Wiener technique and the others is that a method of performing the design computations totally in the time domain can be easily derived for the Wiener technique. To the knowledge of the author, no such derivation has been performed for the Shor or Bryn techniques. This is quite important. The result of the time-domain Wiener design is a set of optimum impulse response functions of a given length. In general, these functions will be different from those resulting from a frequency domain design unless the functions are quite long. Further, for short functions, the Wiener time domain solution will provide better results than the Wiener frequency domain solution. Since it has been indicated that the Wiener frequency domain technique is equivalent to the other techniques, it is tempting to conclude that the Wiener time domain solution will provide generally better results than the other two techniques.

There is a second factor which makes the Wiener time domain solution conceptually more attractive than the Bryn frequency domain solution. This factor has to do with the

difficulty encountered in obtaining good estimates of the cross power density spectra. This is a subject which is well covered in the literature and will not be further discussed here. Let it suffice to say that if all computations are performed in the time domain and the necessity for transforming into the frequency domain is eliminated, then one of the more troublesome aspects of the adaptive system design problem will have been circumvented.

Obviously, as is always the case, there are complications which tend to reduce the attractiveness of the Wiener time domain solution. In particular, the complications have to do with the size of the matrix which must be inverted. For the Bryn technique, the solution requires inverting a great many complex square matrices with the number of complex elements on a side equal to the number of channels being processed. A channel can represent either a single transducer output or a preprocessed subarray of transducers. However, with the Wiener time domain solution, the size of the real matrix to be inverted is equal to the product of the number of channels and the number of filter points in each sample point operator. This matrix is a Toeplitz matrix of submatrices for which a solution (123) is available.

Unfortunately, even though the time-domain matrix is Toeplitz, a time-domain filter design is still considerably slower than a comparable frequency-domain design. Texas Instruments (104) indicates that for comparable 10-channel design problems, the time-domain design requires 333,000 real multiply and add operations whereas the frequency domain design requires only 21,000 complex multiply and add operations. For this example, longer impulse response functions were designed with the frequency domain approach than with the time-domain approach to provide two systems which were of equal effectiveness. From this, it seems reasonable to conclude that, except in rare circumstances, the frequency-domain approach would be more practical to implement than the Wiener time-domain approach.

It seems reasonable to expect that some of the advantages inherent in the Wiener time-domain solution are also possessed by the Widrow wide-band technique. That is, the Widrow wide-band technique optimizes on the basis of a finite number of weights which constitute the elements of a set of finite length impulse response functions. Further, these advantages are obtained without the necessity of either storing or inverting the large matrix required for the Wiener time-domain solution.

As an example, if it were desired to design sample point operators 41 points in length for 50 channels using the Wiener technique, it would be necessary to somehow store 102,500 correlation values. Subsequently it would be necessary to use these 102,500 correlation values to invert a Toeplitz matrix which is 2050 x 2050 in size.

For a comparable Widrow problem, it would be necessary to store 2050 weights, possibly 2050 hydrophone output values, and possibly 2050 signal simulator values, a maximum storage requirement of less than 7000 words. The large matrix inversion would be replaced by an iterative multiply and add process.

Having disposed of the necessity to further consider the Wiener time-domain design approach, the several frequency domain design techniques will be discussed relative to their attractiveness for certain classes of problems.

Two computational operations and their inherent data storage requirements are considered to be most important relative to the implementation of adaptive space processing:

(1) Measurement of correlation functions or frequency domain cross-power functions.

(2) Solution of system equations for design of system parameters.

The Bryn technique requires the solution of the following matrix equation at each frequency of interest:

$$\left\{ H_h(f) \right\} = [N_{ih}(f)]^{-1} \left\{ S^*_{ir}(f) \right\} ,$$

where $H_h(f)$ is the set of complex frequency filters to be applied to each channel,

$N_{ih}(f)$ is the normalized cross-power matrix describing noise,

$S^*_{ir}(f)$ is the conjugate of the normalized cross-power column vector describing signal.

The Wiener frequency domain technique requires the solution of the following matrix equation at each frequency of interest:

$$\left\{ H^*_j(f) \right\} = [S_{ij}(f) + N_{ij}(f)]^{-1} \left\{ S_{ir}(f) \right\} ,$$

where $S_{ij}(f)$ is the cross-power matrix describing signal and the other terms are essentially the same as in the Bryn formulation.

There are three classes of problems which deserve special attention relative to computational requirements. The approach in this paper is to consider first the matrix inversion techniques in an attempt to determine which is most attractive for a particular problem and then to compare the technique thus selected with the iterative techniques as applied to that same problem.

The first problem that will be considered is the use of a small array (less than 50 transducers), with no directional transducer clustering, for the purpose of detecting a signal of

unknown waveform from space organized noise. It will be assumed that continuous monitoring of up to 25 directional beams is required.

The second problem that will be considered is identical to the above problem except that the 50 transducers are arranged in subarray clusters and the clusters are pre-processed in a directional fashion.

The third problem that will be considered is identical to the second problem except that the number of transducers is increased to perhaps 1300.

5.2.1 Small Array Problem With No Clusters

For this problem, there is a significant computational difference between the Wiener frequency domain technique and the Bryn technique. Recall that in the matrix formulation of the Wiener design equations, the signal cross-power functions are made a part of the matrix which must be inverted. This is the characteristic which provides the signal stabilization discussed previously. Consequently, since the signal matrix describing each different signal direction is different from the matrix describing every other signal direction, a different signal plus noise matrix would have to be inverted for each desired directional beam.

Now recall that for the Bryn technique, the signal cross-power functions are not contained in the matrix to be inverted. The directionality of the signal is specified simply by a column vector that is used to multiply the inverse of the noise matrix. Consequently, once that inverse is obtained, the filter weights for many different directional beams can be obtained simply by multiplication of the inverse matrix by many different signal column vectors. Hence, only one matrix inversion

is required for the design of many directional beams. It seems reasonable to conclude then, that for this problem, the Bryn technique is computationally more attractive than the Wiener technique.

If the Bryn technique is used, some method of stabilizing the signal response will probably be required. There is a feedback technique which could be utilized. First the filter weights for one beam could be computed in the normal manner at each frequency. Next the filter weights could be applied to a synthetic signal at each frequency. This would require N complex multiplications at each frequency. Finally a normalization factor could be computed at each frequency which, when applied to all the filter weights, would provide unity signal power in the processed output.

Finally, it is instructive to evaluate the iterative techniques as they apply to small arrays with no directionally sensitive transducer clustering. It will be assumed that a 50-channel problem can be handled by inversion of 50 x 50 complex matrices.

If the Bryn technique is used, at each frequency it will be necessary to measure 1275 complex noise cross-power functions. The remainder of the 2500 cross-power functions in the 50 x 50 matrix are obtained as complex conjugates. In other words, the matrix will be Hermetian and all elements below the diagonal will be the complex conjugates of elements above the diagonal.

Once the inverse of this 50 x 50 matrix is obtained, the filter weights for each of 25 directional beams can be obtained simply by multiplication of the inverse matrix by 25 different column vectors.

Now consider the Shor iterative technique as applied to the same small array problem. In the Shor technique, the design

of the filter weights for each beam must be carried out separately. Considering only one beam at a single frequency, each iteration would require the measurement of 50 complex noise cross-power functions and 50 complex signal cross-power functions or a total of 100 complex cross-power measurements per beam per iteration. For 25 beams, 2500 cross-power measurements would be required for each iteration. It is seen then that for this problem the Bryn technique would require the measurement of substantially fewer cross-power functions than the Shor iterative technique.

It is concluded that for the problem of a small array utilizing no directionally sensitive transducer clustering, processed to provide a large number of directional beam outputs, the Bryn matrix inversion technique is computationally more attractive than either the Wiener matrix inversion technique or the Shor iterative technique. As the number of output beams decreases, the relative attractiveness of the Bryn technique decreases also.

Now consider the use of the Widrow iterative technique for this problem. The storage requirements for a single frequency solution consist of 2500 weights, possibly 50 hydrophone output values and possibly 1250 signal simulator output values, a total of less than 4000 storage locations. In comparison, the Bryn technique would require storage of 2500 complex cross-power values or about 5000 words for the noise matrix alone. In addition to storage of the noise matrix, the Bryn technique would require storage of the history data from which the cross-power values are obtained. This would constitute possibly another 5000 to 10,000 locations. Finally, the Bryn technique would require storage of the same 2500 weights (1250 complex weights) as in the Widrow technique, for a total of possibly 17,500 storage locations.

Unquestionably, the Widrow technique has an advantage over the Bryn technique in terms of required storage. Any statement related to computational difficulty would be premature at this time since the convergence rates for the Widrow technique

when utilized against real sea data are unknown. There are, however, indications that the Widrow technique will prove to be computationally advantageous relative to the Bryn matrix inversion technique, particularly with regard to wide-band operation.

Consideration was not given to the Mermoz matrix inversion technique. This is because the Mermoz technique requires that the exact waveform of the signal be known.

5.2.2 Small Array Problem - Directionally Sensitive Cluster Pre-processing

The next problem to be considered is that of a small array of 50 transducers processed to provide 25 directional beam outputs wherein five directionally sensitive clusters of transducers are formed.

For example, assume the 50 transducers are combined into five subarrays of 10 elements each. Assume each of the 10 small subarrays is processed through conventional beamforming techniques to provide five low-resolution beam outputs. This situation is pictured in Fig. 5.2.

The adaptive space processing problem can then be considered as five separate processors wherein each processor has five inputs. Each adaptive space processor will provide five high resolution adaptive directional outputs where the directions are contained within the low resolution directional beam of the conventional beamformers used as input.

The same considerations apply to each of these parallel space processors as applied to the single space processor discussed in the previous section. Once again, the Bryn matrix inversion technique is computationally more attractive than the Wiener technique; one matrix inversion can be used to provide the complex weighting factors for five adaptive beams with the Bryn technique, and five matrix inversions are required to do the comparable job with the Wiener technique.

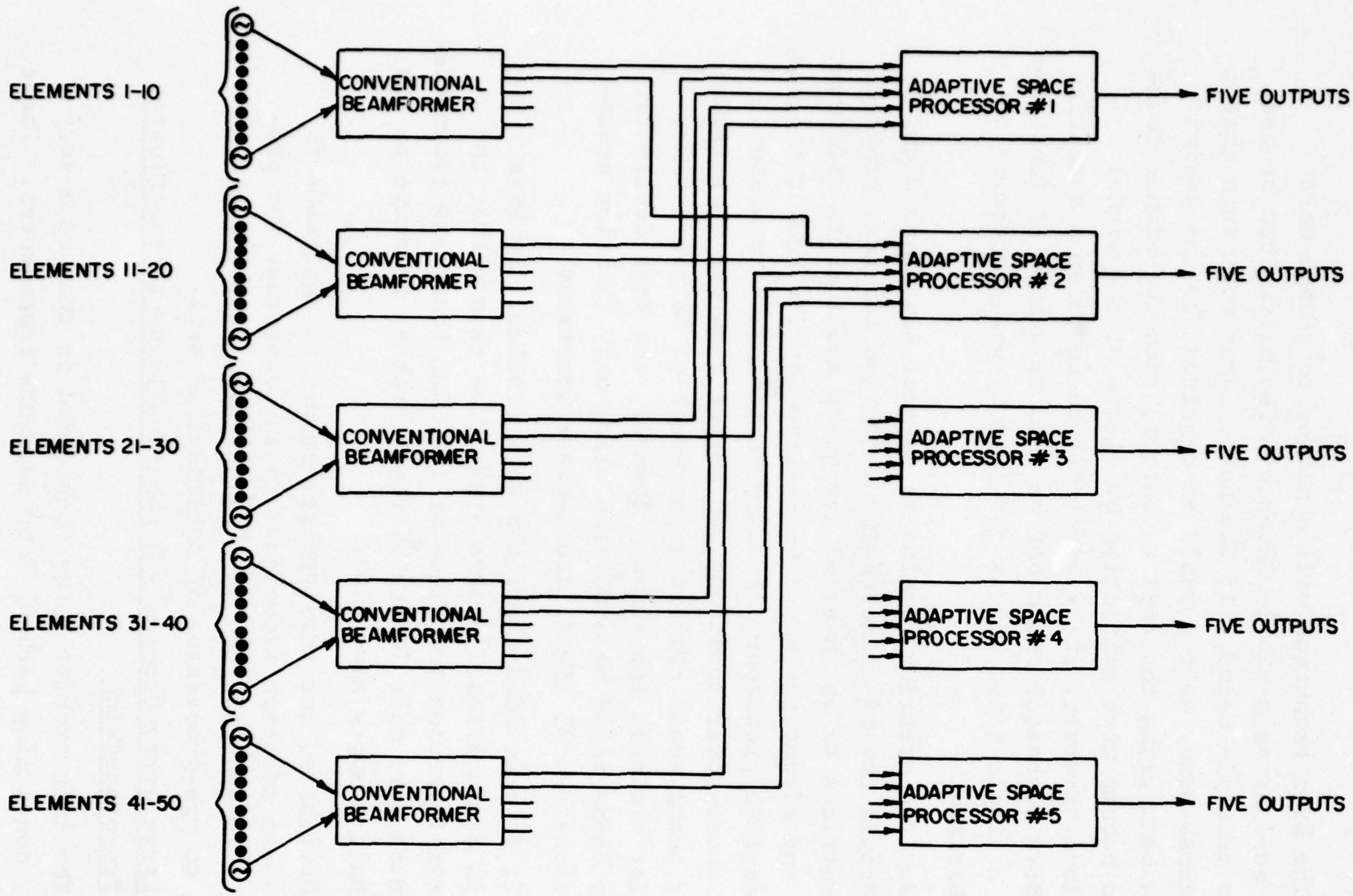

FIG. 5.2— PRE-PROCESSED SMALL ARRAY PROBLEM

TRACOR, INC. DWG. A6-62-59
AUSTIN, TEXAS 10-26-66 BALDWIN/SL

The Bryn technique will continue to remain more attractive so long as a single inverse noise matrix can be used to generate multiple beams. If the noise input were such that a new noise cross-power matrix would be required for the design of each new beam using the Bryn technique, then the Wiener technique would become more attractive by virtue of the signal stabilization. However, it is difficult to envision a situation where the Bryn technique could not be used to design at least two beams from a single noise matrix obtained by pre-processing of smaller subarrays.

Now the question remains as to what advantage might be realized through use of an iterative technique for this problem. Since the matrices to be inverted are quite small it is difficult to see how any advantages would be realized by the Shor technique. For each adaptive processor, it would be necessary to measure 10 complex cross-power functions per output beam per iteration. Five output beams would require measurement of 50 cross-power functions for a single iteration. However, the Bryn design for five output beams could be completed using only 15 noise cross-power functions and 25 signal cross-power functions.

It is concluded that, for this problem, the Bryn technique is computationally more attractive than either the Wiener matrix inversion technique or the Shor iterative technique. Comments which were made in the previous problem relative to the Widrow technique apply here also.

Obviously, for this problem it may be desirable to apply wide-beam adaptive techniques to the subarrays for pre-processing. In general, the conclusions drawn above would be applicable to pre-processing of subarrays as well.

5.2.3 Large Array Problem - Directionally Sensitive Cluster Pre-processing

The last problem to be considered is that of a very large array containing perhaps 1300 separate transducers. There

should be general agrement that with the present state of the art, some sort of subarray cluster processing would be required in order to implement adaptive space processing.

Assume it is desired to provide multiple beam outputs from this array. The problem is very similar to the problem discussed in the previous section. Assume once again that it is desired to provide 25 adaptive beam outputs, and that each of the subarrays is processed using conventional beamforming to provide 25 beam outputs. In other words, each conventional beam from each subarray contributes to only one adaptive beam for the entire array. Assume the 1300 elements are grouped into subarrays containing 26 elements each for a total of 50 subarrays.

This problem can be pictured as 25 separate adaptive space processors each of which acts upon 50 input channels to provide a single adaptive beam output.

With the system organized in this fashion, the previously discussed advantages of the Bryn technique over the Wiener technique disappear. No matter which technique is used, a single matrix inversion would contribute to only one output beam. Thus, it would be necessary to invert a 50 x 50 complex matrix at each frequency for each output beam. Consequently, the Wiener technique would probably be most appropriate due to the inherent signal level stability.

A more reasonable means of operating the system would be to have each subarray provide only 13 beams, with some beam broadening employed if necessary. Then only 13 separate adaptive processors would be required and one noise matrix inversion could be used to design two adaptive output beams. Under this condition, the Bryn technique would be computationally advantageous over the Wiener technique.

Since this is a multiple beam problem, the Shor iterative technique probably would not prove advantageous. For the first

case above, the iterative technique would require the measurement of 2500 complex cross-power functions per iteration to provide 25 beam coverage. The Bryn technique would require the measurement of 31,875 complex cross-power functions.

In order to be comparable on a cross-power measurement basis, the Shor iterative technique would be required to converge in about 13 iterations. It is doubtful if this would be the case.

Now consider another problem associated with the large array. Consider providing, in addition to the 25 preformed adaptive beams mentioned above, two adaptive tracking beams. Let this beam be formed using 200 individual elements as input or 200 sums of very closely spaced elements. Assume that a method for generating signal cross-power models is available based on position and speed information.

To attempt to perform this operation using either the Bryn or Wiener techniques would require inversion of a 200 x 200 matrix each time it were desired to update the beam. Even if this were feasible, it would be necessary to measure 20,100 cross-power functions each update cycle. If a means of using the Shor iterative technique could be implemented, only 400 cross-power measurements per iteration would be required. If the technique would converge in less than 50 iterations, it may prove more attractive than the matrix inversion techniques.

Although some very complicated problems are involved, the example illustrates the conditions under which the Shor iterative technique is attractive. The technique is attractive whenever a large number of transducers is to be processed for the purpose of providing a single or at least a very few output beams. Obviously, for any case where the Shor iterative technique would prove attractive, the Widrow technique would be even more attractive.

5.3 COMPUTATIONAL STABILITY

Another characteristic which differs between the techniques is the computational stability. Although it is difficult to say anything very general in this regard, there are a few points that are worthy of note.

The computational procedure in both the Bryn and Shor techniques will probably tend to become unstable as the percent random noise decreases. This is because for no random noise, it is possible, in some cases at least, to produce an infinite signal-to-noise ratio in the output, and infinity is not mathematically palatable.

The result of having no random noise is to drive the Shor normalized noise correlation indeterminate, and to drive the Bryn matrix singular.

Even though the Bryn noise matrix is singular, the Wiener signal plus noise matrix will not necessarily be singular. Therefore, it is tempting to conclude that the Wiener computations will be generally more stable. However, there is a danger to this conclusion. It is also possible to have a singular Wiener matrix, with a nonsingular Bryn matrix. It is believed, however, that such a matrix would not describe a physically obtainable situation.

5.4 OTHER CHARACTERISTICS

There are other characteristics worthy of analysis which still require study. For example, among the questions for which answers are not yet available are the following: How do the techniques compare when it is desired to control the main beam width to some fixed value? How sensitive are the systems to equalization problems within the array transducers?

BIBLIOGRAPHY

An attempt has been made to provide an extensive bibliography. The following list contains numerous documents in addition to those specifically referenced in the text. This list was made up from two sources:

(a) Documents which are familiar to the author and are considered important.

(b) Documents listed as references by the authors of the documents in (a) above.

A number of these documents have been circulated as internal Navy, company, or university reports and memoranda, but have never been published. For further information pertaining to the availability of documents of this type, it will be necessary to contact the original author.

1. Abbott, Frank, "Simultaneous Multibeam Phase Compensation," NEL Report 642 of 2 December 1955.

2. Angell, J. B., "Multiple Adaptive Logic Functions," Technical Summary Report SEL-65-018, Stanford Electronics Laboratories, Stanford, Calif., (March 1965).

3. Backus, M., J. Burg, D. Baldwin, and E. Bryan, "Wide-Band Extraction of Mantle P-Waves From Ambient Noise," Geophysics, Vol. XXIX, No. 5, pp 672-692 (October 1964).

4. Balakrishnan, A. V., "Estimation and Detection Theory for Multiple Stochastic Processes," J. Math. Anal. and Appl., Vol. I, pp 386-410, (December 1960).

5. Balakrishnan, A. V., "Effect of Linear and Non-Linear Signal Processing on Signal Statistics," J. Res. NBS, Vol. 68D, pp 953-965, (September 1954).

6. Balakrishnan, A. V., "An Adaptive Non-Linear Data Predictor," 1962 Nat'l. Telemetering Conf. Pro.

7. Baldwin, R., "Engineering Aspects of Multichannel Processing Based on Measured Noise Correlations," Proc. IEEE, Vol. 53, No. 12, pp 1884-1889 (Dec. 1965).

8. Bellman, R. E., Adaptive Control: A Guided Tour, Princeton University Press, Princeton, N.J., (1961).

9. Bellman, R. E., Introduction to Matrix Analysis, McGraw-Hill Book Company, Inc., New York, N.Y., (1960).

10. Blackman, R. B., and J. W. Tukey, The Measurement of Power Spectra, New York:Dover, (1959).

11. Brennan, "On the Maximum Signal-to-Noise Realisable from Several Noisy Signals," Proc. IRE (Corresp), Vol. 43, p 1530, (Oct. 1955).

12. Bryn, F., "Optimum Signal Processing of Three-Dimensional Arrays Operating on Gaussian Signals and Noise," J. Acoust. Soc. Am., Vol. 34, No. 3, pp 289-297, (March 1962).

13. Bucy, R. S., "Optimum Finite-Time Filters for a Special Non-Stationary Class of Inputs," JHU/APL Internal Memorandum BBD-600, (1959).

14. Bucy, R. S., "A Matrix Formulation of the Finite-Time Problem," JHU/APL Internal Memorandum BBD-777, (1960).

15. Bucy, R. S., "Combined Range and Speed Gate for White Noise and White Signal Acceleration," JHU/APL Internal Memorandum BBD-811, (1960).

16. Bucy, R. S., "Adaptive Finite-Time Filtering," JHU/APL Internal Memorandum BBD-645, (1959).

17. Burg, J., "Three-Dimensional Filtering With an Array of Seismometers," Geophysics, Vol. XXIX, No. 5, pp 693-713, (October 1964).

18. Carlton, A. G., and J. W. Follin, Jr., "Recent Developments in Fixed and Adaptive Filtering," Proceedings of the Second AGARD Guided Missiles Seminar (Guidance and Control) AGARDograph 21, (September 1956).

19. Claerbout, J. F., "Detection of P Waves from Weak Sources at Great Distances," Geophysics, 29 (2), pp 197-211, (1964).

20. Coddington, E. A., and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill Book Company, Inc., New York, N.Y., (1955).

21. Connolly, J. J., "Fundamental Accuracies in the Simultaneous Measurement of Transverse and Radial Trajectories of Radar Targets," Proc. 7th Nat'l. Conv. Mil. Electronics, Wash., D. C., p 64, (September 1963).

22. Coros, G. M., "Resolution of Pressure in Turbulence," J. Acoust. Soc. Am., 192-199 (1963).

23. Cramer, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, New Jersey, (1946, reprinted 1951), Chap. 24.

24. Davenport, W. B., and W. L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Company, Inc., New York, (1958), Chap. 14.

25. Davisson, Lee D., "A Theory of Adaptive Filtering," IEEE Transactions on Information Theory Vol. IT-12, No. 2, (April 1966).

26. Davisson, L. D., "The Prediction Error of Stationary Gaussian Time Series of Unknown Covariance," IEEE Trans. on Information Theory, Vol. IT-11, pp 527-532, (October 1965).

27. Davisson, L. D., "The Filtering of Time Series With Unknown Signal Statistics," 1965 Nat'l. Electronics Conf. Proc.

28. Davisson, L. D., "Prediction of Time Series From Finite Past," SIAM J., pp 819-826, (October 1965).

29. Dolph, C. L., "A Current Distribution for Broadside Arrays Which Optimizes the Relationship Between Beam Width and Side-Lobe Level," Proc. IRE, Vol. 34, pp 335-348, (June 1946).

30. Doob, J. L., Stochastic Processes, John Wiley & Sons, Inc., New York, N.Y., (1953).

31. Dwork, B. M., "Detection of a Pulse Superimposed on Fluctuation Noise," Proc. IRE, Vol. 38, pp 771-774, (July 1950).

32. Eckart, C., SIO Ref. 52-11, University of California, Marine Physical Laboratory, Scripps Institution of Oceanography (1952).

33. Edelblute, D. J., J. M. Fisk, G. L. Kinnison, New Submarine Sonar/Fire Control System, S23-26, "A Detailed Derivation of the Bryn Equation for Optimal Array Gain, A Comparison with Wiener Processing," NEL Code 2120, U. S. Navy Electronics Laboratory, San Diego, California, (April 1966).

34. Edelblute, D. J., J. M. Fisk, and G. L. Kinnison, Notes on Optimum Signal Detection Theory, 714-222, U. S. Navy Electronics Laboratory, San Diego, California.

35. Faran, J. J., and R. Hills, "Wide-Band Directivity of Receiving Arrays," Harvard Univ. Acoust. Res. Lab. Tech. Mem. 31 (1 May 1953).

36. Faran, J. J., and R. Hills, "Correlation for Signal Reception," Harvard Univ. Acoust. Res. Lab. Tech. Mem. No. 27 (15 Sept. 1952).

37. Faran, J. J., and R. Hills, "The Application of Correlation Techniques to Acoustic Systems," ibid. Tech. Mem. No. 28 (1 Nov. 1952).

38. Fisk, J. M., C. E. Persons, G. L. Kinnison, "New Submarine Sonar/Fire Control System (U), Array Signal Processing - A Review of the Theoretical Equations," NEL Code 2120, (1 Feb. 1966).

39. Gabor, D., "Communication Theory and Cybernetics," Proc. I.R.E., CT-1 (Circuit Theory), 19 (1954).

40. Gabor, D., Wilbry, W. P., and Woodcock, R., "A Universal Non-Linear Filter, Predictor and Simulator Which Optimizes Itself by a Learning Process," Proc. IEE, Part B, 108 (1961).

41. Geronimus, Ya. L., Polynomials Orthogonal on a Circle and Interval, Pergamon Press, London, (1960).

42. Good, I. J., "Speculations on Perceptrons and Other Automata," IBM Research Lecture RC-115, IBM Research Center, Yorktown Heights, New York, (2 June 1959).

43. Goode, B., "An Introduction to the Theory of Sonar Receiving Arrays," Department of the Navy, Code 3130.

44. Grenander, U., and M. Rosenblatt, Statistical Analysis of Stationary Time Series, New York:Wiley, (1957).

45. Hanson, J. E., "Some Notes on the Application of the Calculus of Variations to Smoothing for Finite Time, etc." JHU/APL Internal Memorandum BBD-346, (1957).

46. Heaps, H. S., "General Theory for the Synthesis of Hydrophone Arrays," J. Acoust. Soc. Am. 32, 356-363 (1960).

47. Hiller, E. R., Jr., "Available Source Information in Receiving Apertures," E.E. Thesis Dept. of Elec. Engrg., MIT, Cambridge, Mass. (1956); also Res. Lab. of Electronics MIT, Tech. Rept. No. 3, Contract No. NOrd 14362, (Sept. 1, 1956).

48. Holstrom, C. W., Statistical Theory of Signal Detection, Pergamon Press, N. Y., (1960).

49. Jakowatz, C. V., R. L. Shuey and G. M. White, "Adaptive Waveform Recognition," Fourth International Symposium on Information Theory," ed. by Cherry, Butterworths, London, p 317, (1961).

50. Kahn, "Ratio Square," Proc. IRE, Vol. 42, p 1704, (Nov. 1954).

51. Kailath, T., "Correlation Detection of Signals Perturbed by a Random Channel," IRE Trans. Inform. Theory 6, 361 (June 1960); "Matched Filter Issue," Appendix I.

52. Kalman, R. E., and J. E. Bertram, "Control System Analysis and Design Via the Second Method of Lyapunov. I. Continuous-Time Systems," Journal of Basic Engineering, Trans. ASME, Series D, Vol. 82, pp 371-393, (1960).

53. Kalman, R. E., and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," Journal of Basic Engineering, Trans of ASME, V. 83 D, p 95, (1961).

54. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME, Series D, Journal of Basic Engineering, Vol. 82, pp 35-45, (1960).

55. Kalman, R. E., and J. E. Bertram, "A Unified Approach to the Theory of Sampling Systems," Journal of the Franklin Institute, Vol. 267, pp 405-436, (1959).

56. Kalman, R. E., and R. W. Koepeko, "The Role of Digital Computers in the Dynamic Optimization of Chemical Reactions," Proceedings of the Western Joint Computer Conference, pp 107-116, (1959).

57. Kalman, R. E., "On the General Theory of Control Systems," Proceedings of the First International Congress on Automatic Control, Moscow, USSR, (1960).

58. Kalman, R. E., "Contributions to the Theory of Optimal Control," Proceedings of the Conference on Ordinary Differential Equations, Mexico City, Mexico, (1959); Bol. Soc. Mat. Mex., (1961).

59. Kelly, E. J., L. S. Reed, W. L. Root, "The Detection of Radar Echoes in Noise," Journ. Soc. Indust. & Appl. Math. Vol. 8, Part 1:No. 2, pp 309-341, (June 1960), part D, No. 3, pp 481-501, (September 1960).

60. Kelly, E. J., W. L. Root, "Representations of Vector-Valued Random Processes," Group Rpt. 55-21, Lincoln Lab., 141T, (7 March 1960).

61. Kelly, E. J., Jr., "A Comparison of Seismic Array Processing Schemes," Tech. Note, Lincoln Laboratory, Mass. Institute of Technology, Lexington, Mass., (14 June 1965).

62. Kelly, E. J., Jr., and M. J. Levin, "Signal Parameter Estimation for Seismometer Arrays," Lincoln Laboratory Technical Report 339, (8 Jan. 1964).

63. Kipiniak, W., "Optimum Nonlinear Controllers," Report 7793-R-2, Servomechanisms Lab., M.I.T., (1958).

64. Kullback, S., "Information Theory and Statistics," John Wiley & Sons, New York, N.Y., (1959).

65. Levin, J. J., "On the Matrix Riccati Equation," Trans. American Mathematical Society, Vol. 10, pp 519-524, (1959).

66. Levinson, N., "The Wiener RMS (root mean square) Error Criterion in Filter Design and Prediction," Appendix B of Wiener, N., Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications, pp 129-148, John Wiley and Sons, New York, (1949).

67. Loeve, M., Probability Theory, Van Nostrand and Company, New York, N.Y., Chap. 10, (1955).

68. Lusternik, L.A., and V. J. Sobolev, Elements of Functional Analysis, Gordon & Breach Publishers, New York, pp 149, 283, (1961).

69. Manasse, R., "An Analysis of Angular Accuracies from Radar Measurements," MIT Lincoln Lab., Lexington, Mass., Group Rept. 32-66, (December 1955).

70. Memistor Corporation, Memistor Report, 295 Polaris Avenue, Mountain View, Calif.

71. Mermoz, H., "Adaptive Filtering and Optimal Utilization of an Antenna," Bureau of Ships Translation No. 903, (4 October 1965).

72. Middleton, David, and Herbert L. Groginsky, "Detection of Random Acoustic Signals by Receivers with Distributed Elements: Optimum Receiver Structures for Normal Signal and Noise Fields," Raytheon Tech. Report, Advanced Development Laboratory, Raytheon Company, Wayland, Massachusetts 01778, (Aug. 1963).

73. Middleton, D., An Introduction to Statistical Communication Theory, McGraw-Hill Book Co., Inc., New York, (1960); Topics in Communication Theory, McGraw-Hill Book Co., Inc., New York, (1965).

74. Middleton, D., "On the Detection of Stochastic Signals in Additive Normal Noise," IRE Trans. Inform. Theory 3, 86-149 (June 1957).

75. Middleton, D., "Acoustic Signal Detection by Simple Correlators in the Presence of Non-Gaussian Noise, I. Signal-to-Noise Ratios and Canonical Forms," J. Acoust. Soc. Am. 34, 1598-1609 (1952); "II. Error Probabilities and Canonical Forms for Threshold Operations," Litton Systs., Inc., Waltham, Mass., Tech Rept. TR-64-5-BF (31 Jan 1962). Also, NATO Advan. Study Inst. Symp. Signal Processing with Emphasis on Underwater Acoustics, Centre d'Etude des Phenomenes Aleatoires, Grenoble, 14-26, pp 1-49 (1964).

76. Middleton, D., "On New Classes of Matched Filters and Generalizations of the Matched Filter Concept," IRE Trans. Inform. Theory 6, 349-360, (June 1960).

77. Middleton, D., "On Singular and Nonsingular Optimum (Bayes) Tests for the Detection of Normal Stochastic Signals in Normal Noise," IRE Trans. Inform. Theory 7, 105-113 (Apr. 1961); "Remarks and Revisions of Some Earlier Results in Two Recent Papers," Ibid. 8, 385-387 (Oct. 1962).

78. Middleton, D., D. VanMeter, "Detection and Extraction of Signals in Noise From the Point of View of Statistical Decision Theory," Journ. Soc. Indust. & Appl. Math., part I, Vol. 3, pp 192-253, (December 1955); Part II: Vol. 4, pp 86-119, (June 1956).

79. Ormsby, J., "Design of Numerical Filters with Applications to Missile Data Processing," J. Association for Computing Machinery, pp 440-466, (July 1961).

80. Parzen, E., "An Approach to Empirical Time Series Analysis," J. Res. NBS, Vol. 68D, pp 937-952, (September 1964).

81. Parzen, E., "Statistical Inference on Time Series by Hilbert-Space Methods, I," Tech. Rep. No. 23, Applied Mathematics and Statistics Laboratory, Stanford Univ., (1959).

82. Pugachev, V. S., "General Condition for the Minimum Mean Square Error in a Dynamic System," Avt. i. Telemekh., Vol. 17, pp 289-295, translation, pp 307-314, (1956).

83. Pugachov, V. S., "Teoriyn Sluchainikh Funkstii Ee Primenenio k Zadacham Automaticheskogo Upravieniya," (Theory of Random Functions and Its Application to Automatic Control Problems) (in Russian), second edition, Gostckhizdat, Moscow, (1960).

84. Pugachev, V. S., "A Method for Solving the Basic Integral Equation of Statistical Theory of Optimum Systems in Finite Form" Prikl. Math. Mekh., Vol. 23, pp 3-14 (English translation pp 1-10), (1959).

85. Rice, S. O., "Mathematical Analysis of Random Noise," Bell System Tech. J. 23 (1944); 24 (1945).

86. Robinson, Enders A., "Mathematical Development of Discrete Filters for the Detection of Nuclear Explosion," Jour. Geoph. Res., v. 68, pp 5559-5567, (1963).

87. Robinson, E. A., Random Wavelets and Cybernetic Systems, New York, Charles Griffin and Company, London and Stechert-Hafner, (1962).

88. Rosenblatt, Frank, "The Perceptron. A Theory of Statistical Separability in Cognitive Systems," Cornell Aeronautical Laboratory, Inc., Report No. VG-1196-G-1 (January 1958).

89. Rozonoer, L. I., "L. S. Pontryagin's Maximum Principle in the Theory of Optimum Systems, I," Avt. i Telemekh., Vol. 20, pp 1320-1324, (1959).

90. Schmid, C., "Digital Computer Phase Techniques for Determining Noise Source Locations," Paper presented at 68th Meeting of the Acoustical Society of America, in Austin, Texas, on October 24, 1964, p 3.

91. Schneider, W. A., K. L. Larner, J. P. Burg, and M. M. Backus, "A New Data Processing Technique for the Elimination of Ghost Arrivals on Reflection Seismograms," Geophysics, 29(5), 783-805, (1961).

92. Selfridge, Oliver G., and Ulric Neisser, "Pattern Recognition," Scientific American, 203, 2, 60, (August 1960).

93. Shinbrot, M., "Optimization of Time-Varying Linear Systems With Nonstationary Inputs," Trans. ASME, Vol. 80, pp 457-462, (1958).

94. Shor, S. W. W., "Adaptive Technique to Discriminate Against Coherent Noise in a Narrow Band System," J. Acoust. Soc. Am., Vol. 39, No. 1, pp 74-78 (January 1966).

95. Simpson, S. M., E. A. Robinson, R. A. Wiggins, and C. I. Wunsch, "Studies in Optimum Filtering of Single and Multiple Stochastic Processes," Sci. Rept. 7, Contract AF 19(604)7378, Massachusetts Institute of Technology, Cambridge, (1963).

96. Steeg, C. W., "A Time-Domain Synthesis for Optimum Extrapolators," Trans. IRE, Prof. Group on Automatic Control, pp 32-41, (Nov. 1957).

97. Stocklin, P. L., "Adaptive Filtering in Underwater Sound," U. S. Office of Naval Research, London, England.

98. Stocklin, P. L., "Limits of Measurement of Acoustic Wave Fields," Ph.D. Thesis, University of Conn., Dept of Physics, (1962). Available for purchase from University Microfilms, Ann Arbor, Mich., 62-4398.

99. Stocklin, P. L., "Space-time Sampling and Likelihood Ratio Processing in Acoustic Pressure Fields," Electronic and Radio Engineer (Journ. Brit. IRE) Vol. 26, No. 1, pp 79-91, (July 1963). Presented at the Symposium on Sonar Systems, Birmingham, U.K., pp 9-11, (July 1962).

100. Stocklin, P. L., "Limits of Measurement of Acoustic Wave Fields," Ph.D. thesis, Univ. Conn. (1962).

101. Stocklin, P. L., "Adaptive Filtering in Passive Detection," paper in preparation, Office of Naval Research, London.

102. Systems Theory Laboratory, Stanford Electronic Laboratories, Stanford, California, "Proposal for Research on Adaptive Hydrophone Arrays."

103. Szego, G., "Orthogonal Polynomials," American Mathematical Society, New York, (1959).

104. Texas Instruments, Inc., Final Report, NEL Contract N123(953)56365A, Document Registration No. 20491.

105. Texas Instruments, Inc., Seismometer Array and Data Processing System: Final Report, Phase I, AFTAC Project No. VT/077, Contract AF 33(600)-41840, (1961).

106. Thomas, J. B., L. A. Zadoh, "Note on an Integral Equation Occurring in the Prediction, Detection and Analysis of Multiple Time Series," IRE Trans. Info. Theory, Vol. IT-7, No. 2, pp 118-120,(April 1961).

107. Thomas, J. B., "On the Statistical Detection Problem for Multiple Signals," IRE Trans. on Info. Theory, Vol. IT-8, No. 4, pp 274-280, (July 1962).

108. Turin, G. L., "On Optimal Diversity Reception," IRE Trans. on Info. Theory, Vol. IT-7, No. 3, pp 154-166; (July 1961).

109. Urkowitz, H., "The Accuracy of Maximum Likelihood Estimates of Angle in Radar and Sonar," IEEE Trans. Military Electronics, Vol. MIL-8, No. 1, (January 1964). Also as Report PR-315, Philco Corp. Scientific Lab., Blue Bell, Pa., (September 1963).

110. Urkowitz, H., "On Detection and Estimation of Wave Fields for Surveillance," IEEE Transactions on Military Electronics, Vol. MIL-9, pp 44-56, (Jan. 1965).

111. Urkowitz, Harry, "Space-Time Detection and Estimation," An Address before the Graduate Seminar, Princeton University, Department of Electrical Engineering, (February 13, 1964).

112. Vanderkulk, W., "Optimum Processing for Acoustic Arrays," Journal Brit. I.R.E., p 286, (Oct. 1963).

113. Vanderkulk, W., "Optimum Processing of Active Acoustic Arrays (U)," Section I-D, Vol. III, Final Report, Institute for Defense Analyses Sonar Signal Processing Summer Study, Naval War College, Newport, R.I., (July 1963).

114. Weaver, C. S., "Adaptive Communication Filtering," IEEE Trans. on Information Theory, Vol. IT-8, pp S169-S178, (September 1962).

115. Welch, P. D., "A Direct Method of Power Spectrum Estimation," IBM J. Res. Developm., pp 141-157, (April 1961).

116. Westcott, J. H., "Design of Multivariable Optimum Filters," Trans. ASME, Vol. 80, pp 463-467, (1958).

117. Widrow, B., "An Adaptive 'Adaline' Neuron Using Chemical 'Memistors'," Technical Report No. 1553-2, Stanford Electronics Laboratories, Stanford, Calif., (October 1960).

118. Widrow, B., et al, "Practical Applications for Adaptive Data Processing Systems," IEEE WESCON Reprint II, 4 (August 1963).

119. Widrow, B., "Adaptive Sampled-Data Systems," class notes, Department of Electrical Engineering, Stanford University, Stanford, Calif., (May 1964).

120. Widrow, B., "An Introduction to Adaptive Threshold Logic Decision-Making Systems," class notes, Department of Electrical Engineering, Stanford University, Stanford, Calif., (May 1965).

121. Widrow, B., and Goode, B., Monthly Reports, Contract NObsr-95038.

122. Wiener, N., Extrapolation, Interpolation, and Smoothing of Stationary Time Series With Engineering Applications, New York, John Wiley and Sons, Inc.

123. Wiggins, R., and E. Robinson, "Recursive Solution to the Multichannel Filtering Problem," J. Geophysical Research, Vol. 70, No. 8, pp 1885-1890, (April 15, 1965).

124. Wolf, J. K., "On the Detection and Estimation Problem for Multiple Nonstationary Random Processes," Ph.D. Thesis, Princeton U., Dept. of EE, (1960). Available for purchase from University Microfilms, Ann Arbor, Mich. Mic. 60-5074.

125. Wong, E., "Vector Stochastic Processes in Problems of Communication Theory," Ph.D. Thesis, Princeton U., Dept. of EE., (1959). Available for purchase from University Microfilms, Ann Arbor, Mich., Mic. 59-5234.

126. Yaglom, A. M., "Vvedenie v Teoriya Statsionarnikh Sluchainikh Funktsii," (Introduction to the Theory of Stationary Random Processes) (in Russian) Ups. Fiz. Nauk., Vol. 7, (1951); German translation edited by H. Goring, Akademio Verlag, Berlin, (1959).

127. Yaglom, A. M., "On the Equivalence and Perpendicularity of Two Gaussian Probability Measures in Function Space," in Time Series Analysis, M. Rosenblatt, Ed. John Wiley & Sons, Inc., New York, (1963), Chap. 22.

128. Yaglom, A. M., An Introduction to the Theory of Stationary Random Functions, Prentice-Hall Publ. Co., Inc., Englewood Cliffs, N. J., (1960), Chap. 3, Secs. 15-17.

APPENDIX A

APPENDIX A
COMPUTATIONAL EXAMPLES FOR WIDROW-GOODE TECHNIQUE

In order to develop some feeling for the physical significance of the Widrow-Goode iterative technique, it was instructive to apply the technique to a very simple problem and observe the variations of certain parameters as a function of the number of iterations.

A particularly simple problem was chosen for which the optimum solution could be obtained by inspection. The problem is illustrated in Figure A-1. Three omnidirectional hydrophones were assumed to be arranged equally spaced on a line perpendicular to the desired "look" direction, or axis, for signal. The two outer hydrophones were assumed to be averaged together. A single frequency noise source was assumed to lie at a bearing angle of 60° off the look axis, as shown. Hence a noise attenuation of 6 dB was obtained in the output at the two element "cluster." It was assumed that the signal was also single frequency, and that ambient and circuit noise was non-existent. With this configuration, the optimum solution was obtained whenever W1 was equal to -1 and W2 was equal to +2. This set of weights provided an infinite signal-to-noise ratio in the processor output. Using this configuration it was not necessary to provide phase shifts other than zero or 180 degrees. Hence, the time delay networks could be omitted.

It was assumed that the iteration rate was an integral submultiple of the signal and noise frequency. Hence, each time a measurement was made of the single frequency signal amplitude or noise amplitude, the same value was obtained. It was assumed that measurements were made at times corresponding to the peaks of the signal and noise sinusoids. Signal measurements always yielded unity on both channels. Noise measurements always yielded unity on channel No. 1 and 0.5 on channel No. 2.

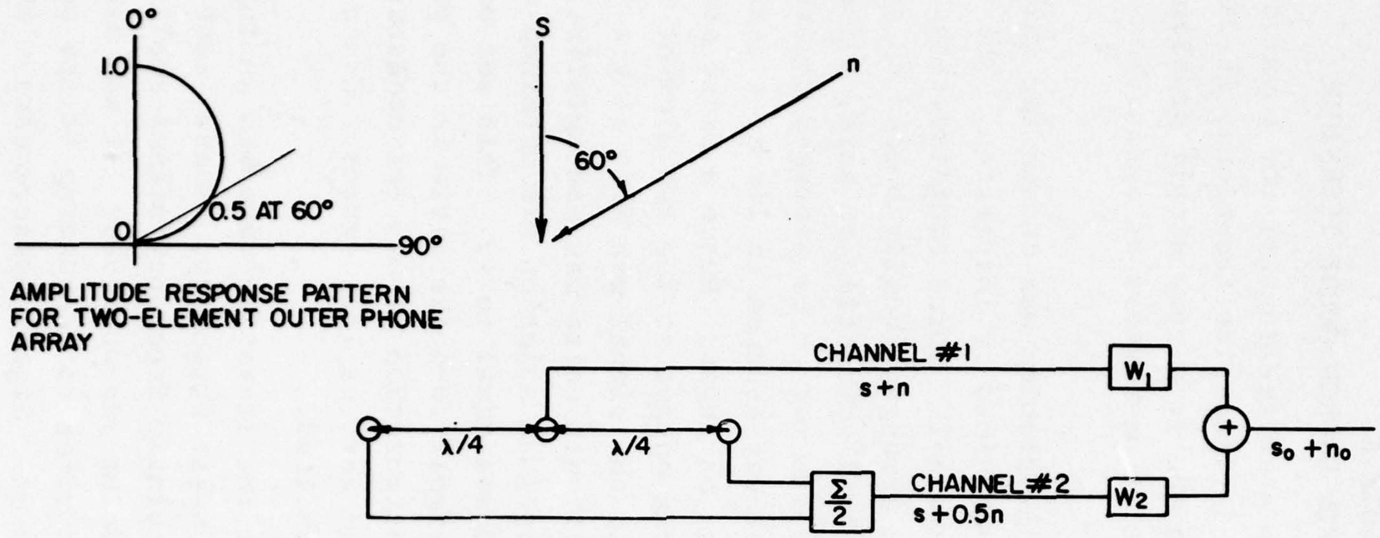

FIG. A-1 — ARRAY PROBLEM USED TO ILLUSTRATE WIDROW-GOODE NARROW-BAND TECHNIQUE

TRACOR, INC. DWG. A6-62-60
AUSTIN, TEXAS 10-26-66 BALDWIN/SL

The signal error was defined to be

$$ES = 1.0 - (W1 + W2),$$

and the noise error was defined to be

$$EN = -(W1 + 0.5\ W2).$$

The output signal amplitude was defined to be

$$SIGAMP = W1 + W2.$$

The output signal to noise ratio was defined to be

$$SNRAT = -SIGAMP/EN.$$

During mode II, the weighting functions were adjusted according to

$$W1 = W1 + 0.1\ EN,$$

$$W2 = W2 + 0.05\ EN.$$

Alternately, during mode I, the weighting factors were adjusted by

$$W1 = W1 + 0.1\ ES,$$

$$W2 = W2 + 0.1\ ES.$$

It was assumed that W1 and W2 were each equal to 0.5 at the beginning of the iterative process.

Experimental results are shown in Figure A-2. For this experiment, the time spent in mode I was equal to the time spent in mode II. Several interesting parameters are shown plotted in Figure A-2 as a function of the number of iterations.

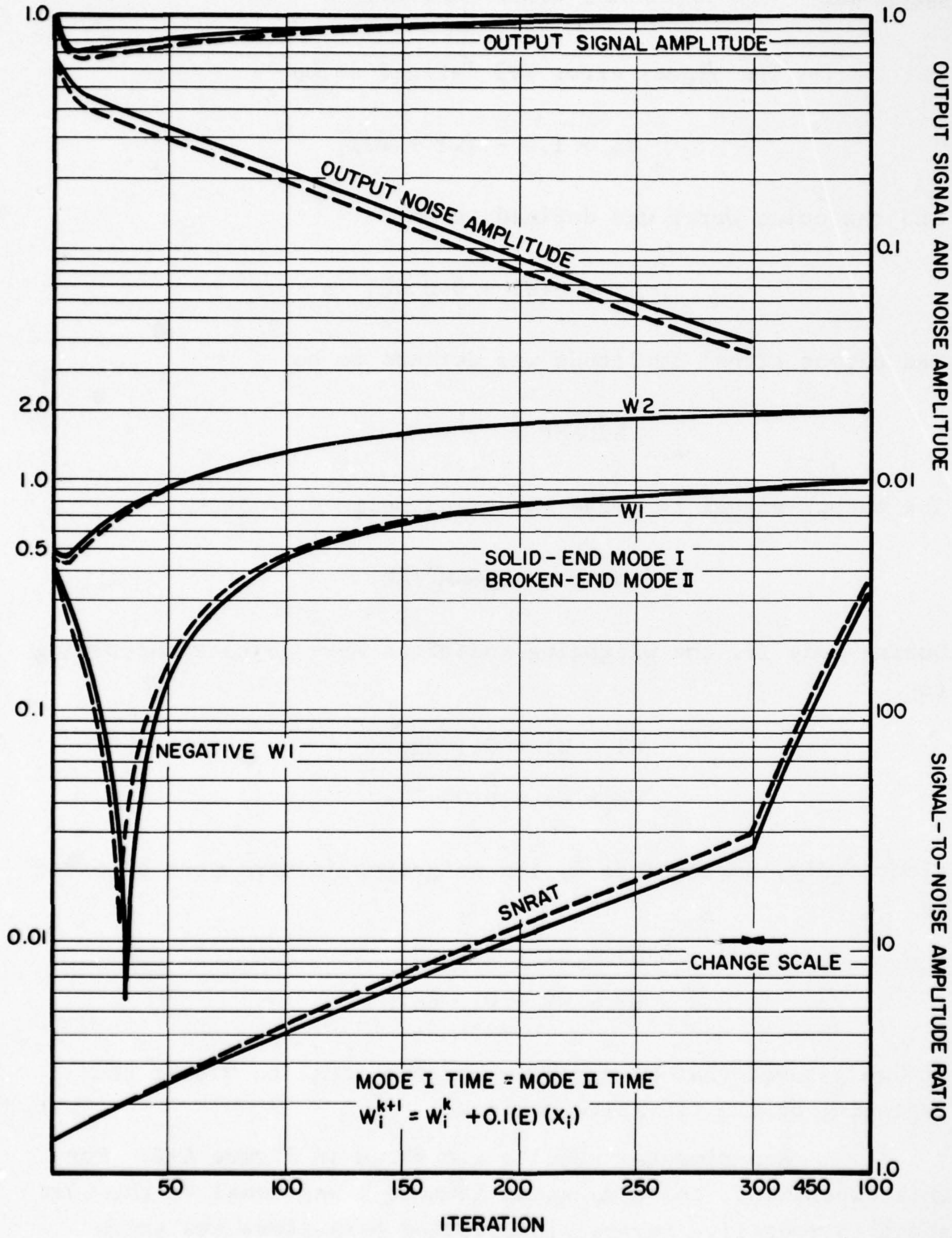

FIG. A-2 — TWO-CHANNEL WIDROW-GOODE EXPERIMENT

TRACOR, INC. DWG. A6-62-53
AUSTIN, TEXAS 10/17/66 BALDWIN/SL

The system was switched alternately between the two modes. One iteration included sequential operation in both modes. For each iteration, the results at the end of mode I and at the end of mode II are shown plotted. The values of the parameters can be thought of as alternately switching back and forth between the broken and solid lines during each iteration.

The curves are particularly well behaved, with the values of W1 and W2 asymptotically approaching the optimum values of -1 and 2 respectively. Of particular interest to the author was the manner in which the value of the output noise amplitude varied as the system was switched from mode I to mode II. Whenever the switch was made and the system was taught to preserve signals, the noise in the output increased by 1.15 dB. Early in the iterative process, training the system to reject noise caused a comparable change in the signal response. Hence, the S/N following mode II was little different from the S/N following mode I. By the time 150 iterations had been completed, training the system to reject noise had little effect on the signal response but training the system to preserve signals consistently reduced the noise rejection by 1.15 dB. Hence, the S/N at the end of modes I and II began to spread.

Another interesting aspect of the curves is the logarithmic (or dB) linearity of the output noise amplitude following the initial fluctuations. Logarithmic linearity of the output noise with a negative slope indicates that the output noise was reduced by the same factor each iteration. The noise at the end of each mode II was consistently 99.2 percent of the noise at the end of the previous mode II.

The space filtering problem can be considered as a problem of finding a weight vector which is orthogonal to the noise vector. However, the weight vector must be such that multiplication by the signal vector will yield the desired result.

This concept is illustrated pictorially in Figure A-3 which shows the signal vector, the noise vector, the optimum weight vector, and the locus of the actual weight vector obtained by iteration. An axis at orthogonality is shown by one of the broken lines. Any weight vector which lies on this axis would be orthogonal to the noise vector, and hence would satisfy the noise requirement.

There are an infinite number of vectors which will satisfy the signal constraint. The locus of these vectors is shown by a broken line. However, there is only one vector which will satisfy both the signal constraint and the noise constraint simultaneously. That is the vector defined by the intersection of the axis of orthogonality and the locus of all vectors which satisfy the signal constraint.

The process of training the system to reject noise can, therefore, be thought of as adjusting the direction of the weight vector toward the axis of orthogonality. The alternative process of training the system to preserve signal can be thought of as adjusting the weight vector so as to provide the proper product between the weight vector and the signal vector.

As the noise vector and the signal vector become more nearly equal, the length of the weight vector would be required to increase in order to preserve signal. This is because the angle between the signal vector and the axis of orthogonality would approach 90°. The cosine of the angle would decrease, and hence the product

$$SW \cos \theta_{SW}$$

would also decrease unless W increased proportionately. This leads to the "super gain" array concept encountered in radar applications several years ago.

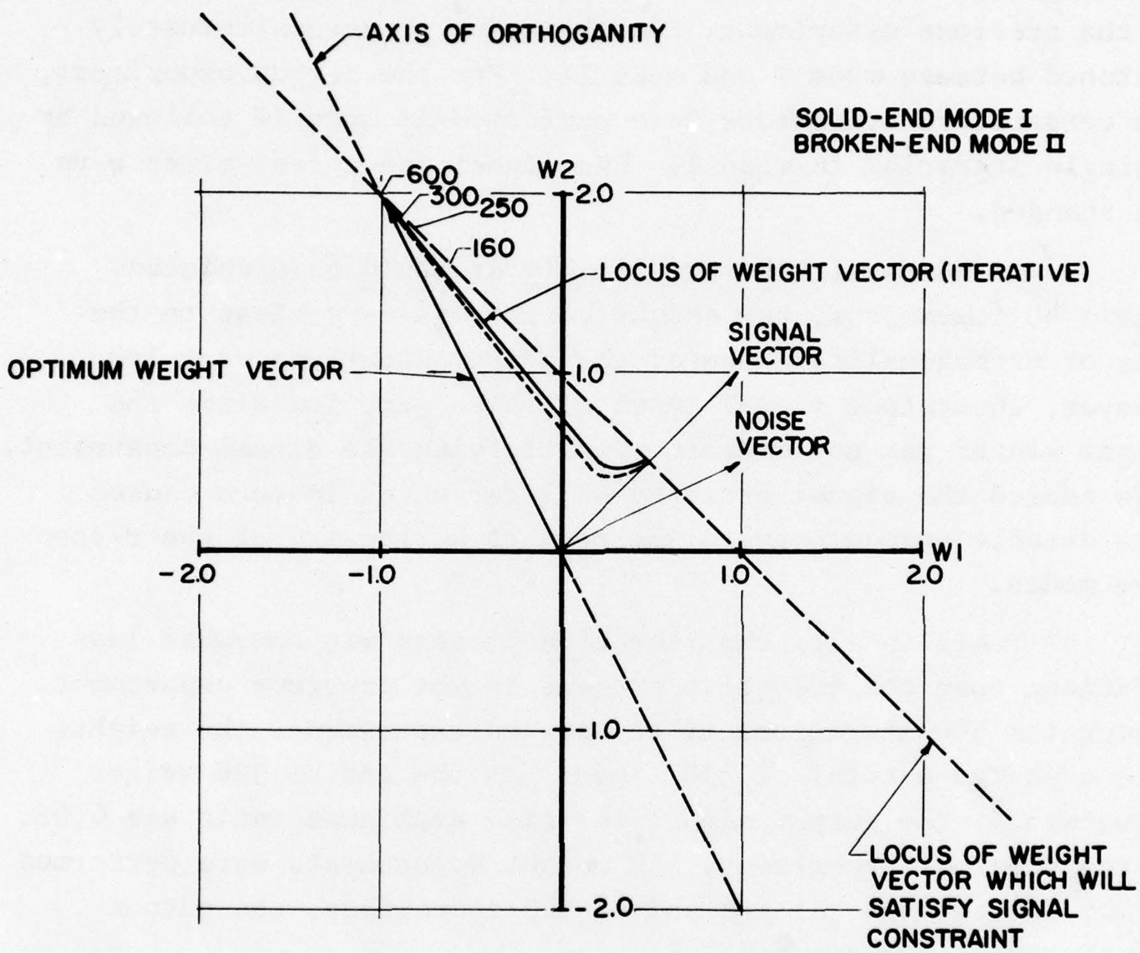

FIG. A-3 — LOCUS OF WEIGHT VECTOR IN TWO-CHANNEL WIDROW-GOODE EXPERIMENT. REF. FIG. A-2

TRACOR, INC. DWG. A6-62-54
AUSTIN, TEXAS 10/17/66 BALDWIN/SL

Figures A-4 and A-5 show the results of a second experiment. The only difference between this experiment and the previous experiment was the percentage of time spent in mode I. In the previous experiment, the iterative process alternately switched between mode I and mode II. For the second experiment, ten consecutive iterations were performed in mode II followed by a single iteration in mode I. The signal and noise values were not changed.

The results were generally as would be predicted. Within 40 iterations, the weight vector was very close to the axis of orthogonality. Hence, the noise output was very low. However, the output signal level was also very low since the weight vector was not capable of satisfying the signal constraint. This caused the signal error to be large which in turn caused considerable spread between the results at the end of the respective modes.

All in all, the iterative process was somewhat less efficient than the iterative process in the previous experiment. During the 300 iterations of the second experiment, the weights were adjusted a total of 330 times. At the end of 330 weight adjustments, the output signal-to-noise amplitude ratio was 6.05. In the previous experiment, 330 weight adjustments were performed in 165 iterations. At the end of 165 iterations, the output signal-to-noise ratio was 7.5.

The remarks in the previous paragraph were based on the assumption that the iterative process should be ended in a mode I operation. If, on the other hand, signal response were not a major consideration and the output signal-to-noise ratio were of prime importance, then the iterative process in the second experiment would be most efficient. At the end of 329 weight changes, the output signal-to-noise ratio was 18.94. In the previous experiment, 255 total iterations or 510 weight changes were required to produce an output signal-to-noise ratio this large.

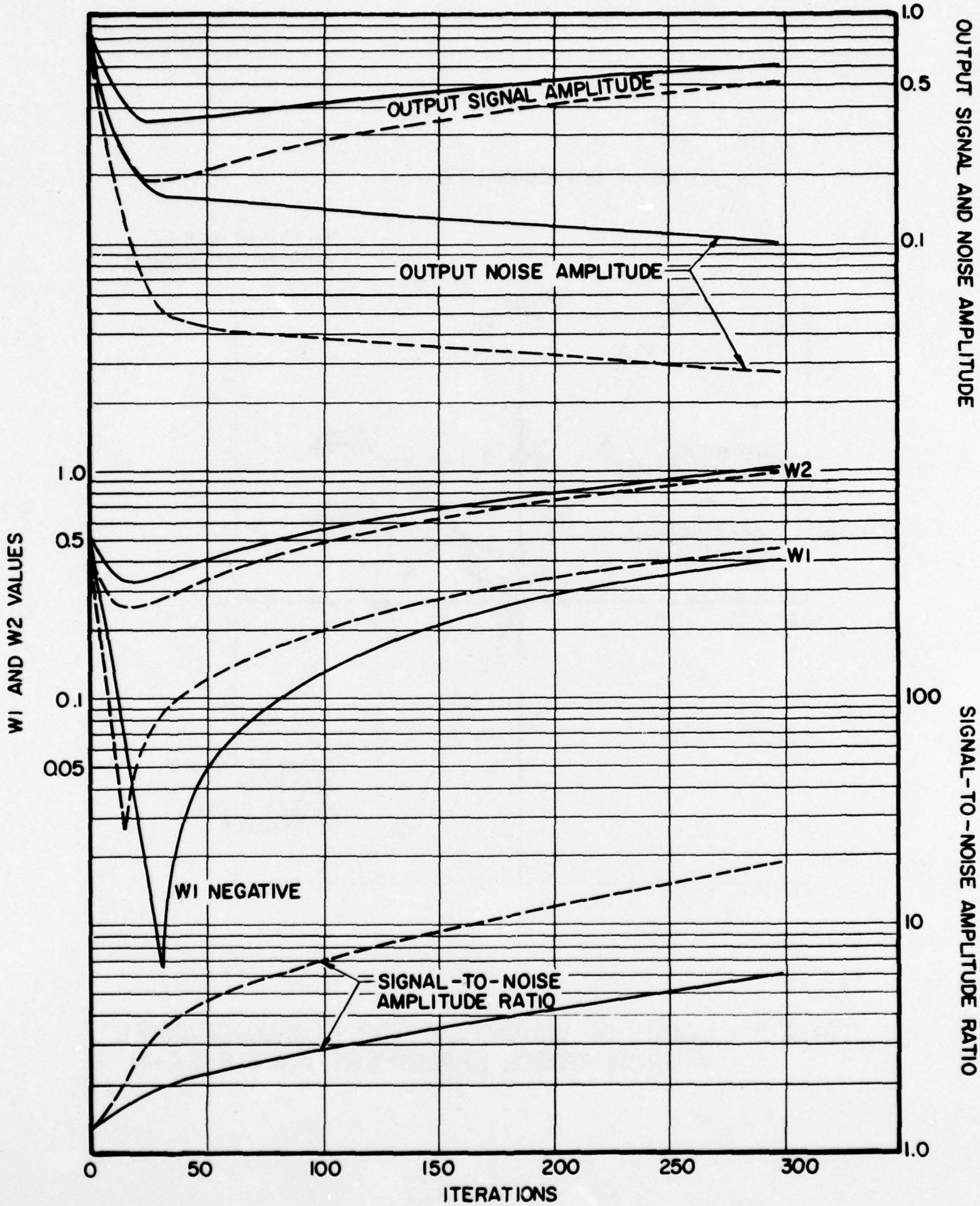

FIG. A-4 -TWO-CHANNEL WIDROW-GOODE EXPERIMENT

TRACOR, INC. DWG. A6-62-63
AUSTIN, TEXAS 10-26-66 BALDWIN/SL

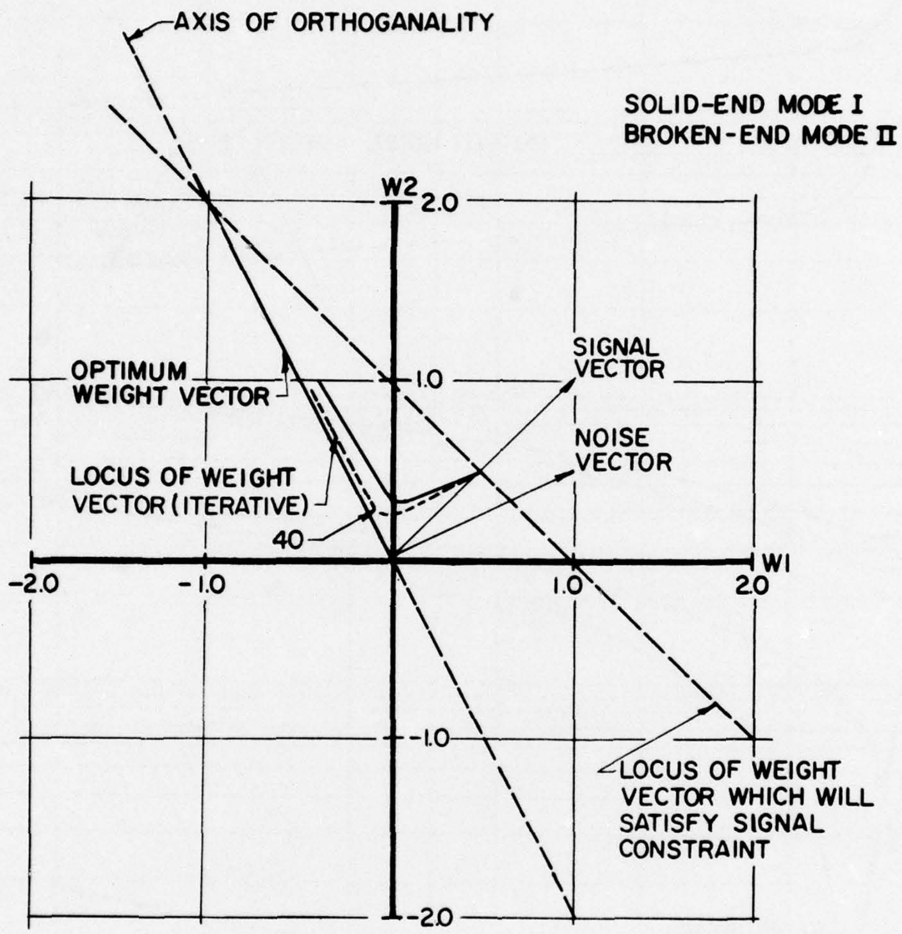

FIG. A-5 — LOCUS OF WEIGHT VECTOR IN TWO-CHANNEL WIDROW-GOODE EXPERIMENT. (REF. FIG. A-4)

TRACOR, INC. DWG. A6-62-61
AUSTIN, TEXAS 10-26-66 BALDWIN/SL

Those familiar with the Wiener technique will recognize this tradeoff between signal preservation and signal-to-noise improvement as a common occurrence in space processing. Quite often, if a spatial processing system is allowed to produce signal distortion, it will also produce more processing gain than a similar system constrained to minimize signal distortion. In fact, techniques such as the Mermoz technique which are designed to maximize the signal-to-noise ratio may be thought of as producing signal distortion in an optimum manner.

Figures A-6 and A-7 show the results of a third experiment. This experiment differed from the previous experiment in one respect. For the third experiment, whenever the time came to adjust the weights so as to satisfy the signal requirement, the weight vector was simply lengthened without changing its direction. Hence, the signal response was improved with no degradation in the signal-to-noise ratio.

The top illustration in Figure A-6 shows the manner in which the weight vector was changed through the iterative process of experiment number 2. Ten iterations were used to position the weight vector toward the axis of orthogonality. Following this, one iteration was used to position the weight vector toward the locus of all possible vectors which would properly preserve signal. It can be seen that whenever the weight vector was repositioned so as to preserve signal, it was also changed in direction. Hence, a decrease in S/N occurred whenever the system was taught to preserve signals.

The bottom illustration in Figure A-6 shows the manner in which the weight vector was changed through the iterative process of experiment number 3. The additive iteration process used to adjust the weights for signal preservation in experiment number 2 was replaced by a multiplicative adjustment process in experiment number 3. Hence, the operation

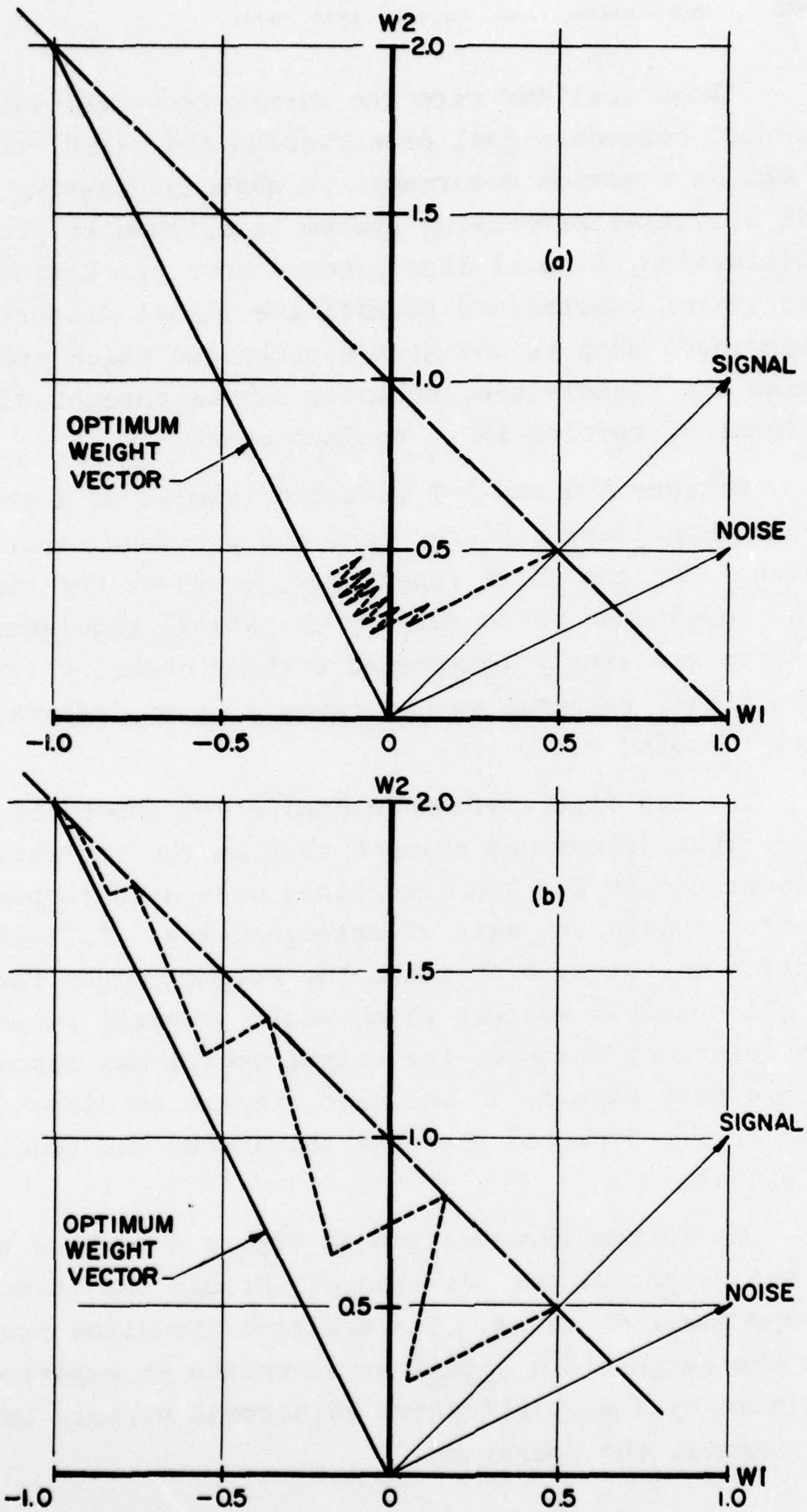

FIG. A-6 —WEIGHT VECTOR ITERATIONS FOR (a) EXPERIMENT NUMBER 2, (b) EXPERIMENT NUMBER 3

TRACOR, INC. DWG. A6-62-62
AUSTIN, TEXAS 10-26-66 BALDWIN/SL

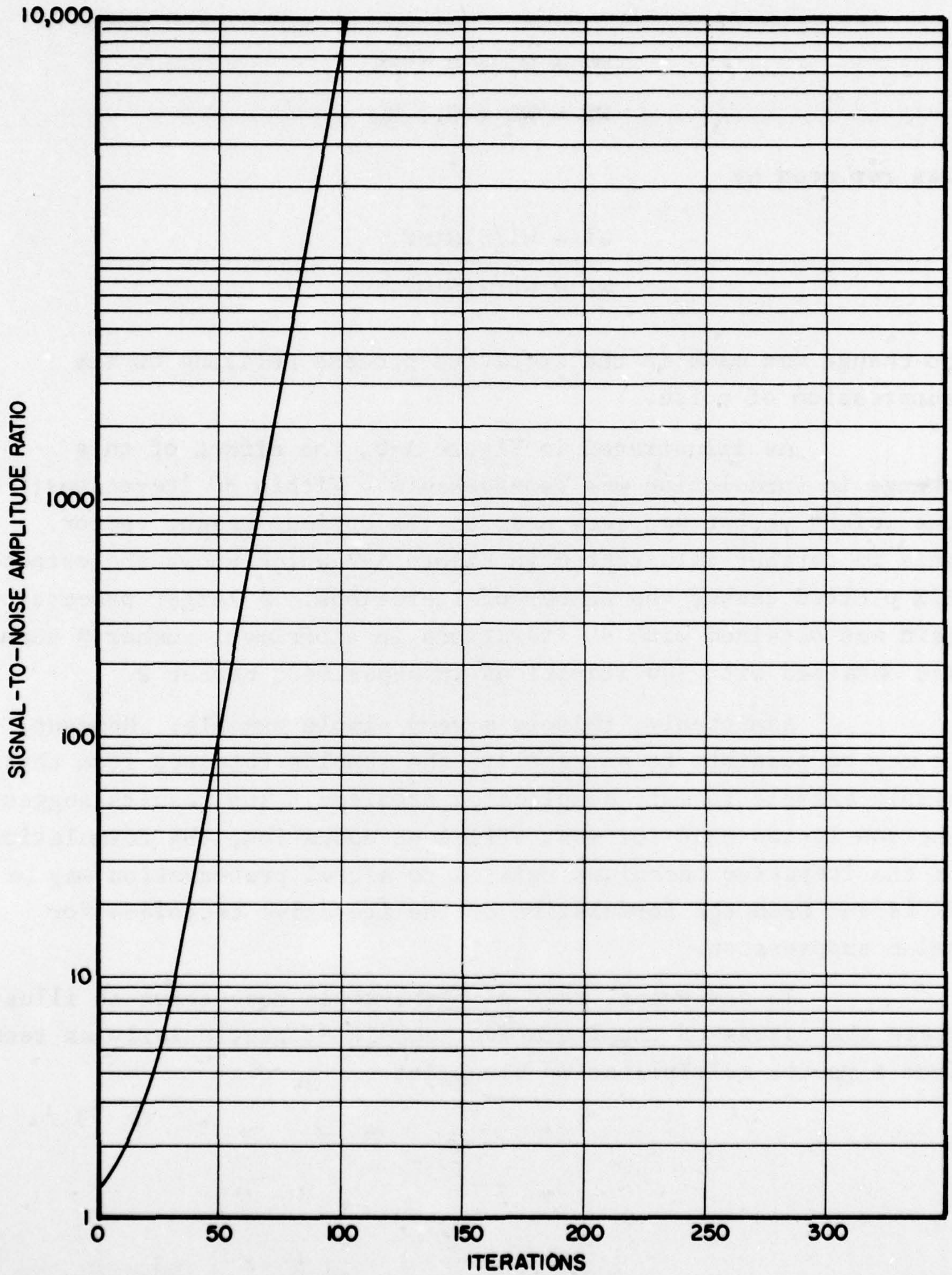

FIG. A-7—OUTPUT S/N FOR EXPERIMENT NUMBER 3

TRACOR, INC. DWG. A6-82-64
AUSTIN, TEXAS 10-26-66 BALDWIN/SL

W1 = W1 + 0.1 ES

W2 = W2 + 0.1 ES

was replaced by

W1 = W1/SIGAMP

W2 = W2/SIGAMP.

No change was made in the iterative process relating to the suppression of noise.

As illustrated in Figure A-6, the effect of this change in formulation was considerable. Within 50 iterations, the weight vector was very near to the optimum weight vector. This is further illustrated in Figure A-7 which shows the output S/N plotted versus the number of iterations. A larger processing gain was obtained with 40 iterations in experiment number 3 than was obtained with 300 iterations in experiment number 2.

Admittedly, this is a very simple example. However, it may be possible to extrapolate the results obtained from this simple example to more complicated problems. The results suggest the conclusion that for most efficient operation, the formulation of the iterative technique related to signal preservation may be different from the formulation of the iterative technique for noise suppression.

In any event, this simple example does serve to illustrate the nature of the iterative technique, particularly as seen from a vector multiplication viewpoint.

APPENDIX B

APPENDIX B
OUTPUT SIGNAL LEVEL CONSIDERATIONS FOR WIENER AND BRYN PROCESSORS

A very simple problem was worked in order to illustrate the output signal level characteristics of the Wiener and Bryn processors.

The problem is as illustrated in Fig. B-1 below. Three hydrophones are arranged in a line array with a spacing of 1/2 wavelength. The two outer phones are averaged to provide channel number one for the processor. The center phone provides channel number two for the processor.

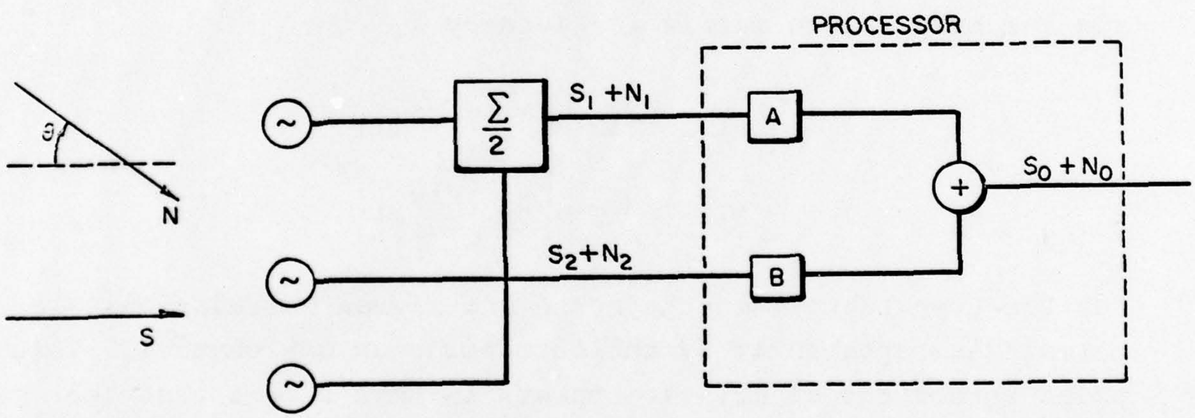

FIG. B.1 — ARRAY PROBLEM USED TO ILLUSTRATE SYSTEM SIGNAL RESPONSE

FIGURE B-1. ARRAY PROBLEM USED TO ILLUSTRATE SYSTEM SIGNAL RESPONSE

The array is excited by signals arriving broadside to the array and by noise arriving with some angle θ as shown. With the two outer phones added in this manner, all cross-power functions must be purely real. Hence, A and B can be real weighting factors.

Appropriate calculations were made so that the output signal level, output noise level, and the output signal-to-noise ratio for the two techniques as a function of the angle θ could be compared.

The Wiener design matrix is given by

$$\begin{bmatrix} S_{11} + N_{11} & S_{12} + N_{12} \\ S_{21} + N_{21} & S_{22} + N_{22} \end{bmatrix} \begin{Bmatrix} A \\ B \end{Bmatrix} = \begin{Bmatrix} S_{11} \\ S_{21} \end{Bmatrix}$$

and the Bryn design matrix is given by

$$\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{Bmatrix} A \\ B \end{Bmatrix} = \begin{Bmatrix} S_{11} \\ S_{21} \end{Bmatrix}$$

For the Bryn technique, the noise and signal correlations are normalized, apparently by the auto power on one channel. This point is not too clear. It appears in Bryn's text that the assumption is made that the average noise power on all channels must be equal and the average signal power on all channels must be equal. This is an unduly severe restriction which would eliminate the technique from practical consideration at the outset. Therefore, Bryn's normalized correlation has been interpreted to be normalized simply by the average power on a single channel.

With the Wiener technique, no such assumption is necessary. However, in order to provide a basis for comparison,

the following problem was worked with the signal and the noise correlation matrices normalized on channel number 2.

Four cases were considered with θ equal to 90°, 9.6°, 4.15° and 0° respectively. Pertinent results are contained in Table number I.

The effect on the output signal power level of varying θ is clearly illustrated. As θ varies from 90° to 0°, the output power for the Bryn technique varies from 40,000 to 1. This would require a 46 dB dynamic range for the equipment operating on the output signal. The output signal power for the Wiener technique varies between 0.989 and 0.25, a change of only 5.96 dB. Further, the output signal power for the Wiener processor could be readily stabilized to within 3 dB with no appreciable change in the space processing gain. No straightforward method of stabilizing the output signal level from the Bryn processor in the matrix inversion is apparent at this time. However, a feedback type of stabilization can be implemented.

TABLE I

		WIENER					BRYN				
θ	NOISE MATRIX	A	B	NOISE OUT	SIG OUT	S/N OUT	A	B	NOISE OUT	SIG OUT	S/N OUT
90°	$\begin{bmatrix} 1.01 & -1.0 \\ -1.0 & 1.01 \end{bmatrix}$	0.498	0.498	0.0049	0.989	200	100	100	200	40,000	200
9.6°	$\begin{bmatrix} 0.26 & 0.50 \\ 0.50 & 1.01 \end{bmatrix}$	1.801	-0.848	0.04	0.909	22.7	40.5	-19.05	21.45	461	21.45
4.15°	$\begin{bmatrix} 0.82 & 0.90 \\ 0.90 & 1.01 \end{bmatrix}$	2.29	-1.66		0.4		6.05	-4.39		70.8	
0°	$\begin{bmatrix} 1.01 & 1.00 \\ 1.00 & 1.01 \end{bmatrix}$	0.25	0.25	0.25	0.25	1.0	0.5	0.5	1.0	1.0	1.0

NOTES:

(1) The signal matrix is given by $\begin{bmatrix} 1.0 & 1.0 \\ 1.0 & 1.0 \end{bmatrix}$ in all cases.

(2) The output noise, output signal, and output signal-to-noise ratio are given in power.

(3) One percent random noise was assumed.